Census 2000 Testing,
Experimentation,
and Evaluation Program
October 25, 2002

Response Mode and Incentive Experiment for Census 2000

FINAL REPORT

This research paper reports the results of research and analysis undertaken by the U.S. Census Bureau. It is part of a broad program, the Census 2000 Testing, Experimentation, and Evaluation (TXE) Program, designed to assess Census 2000 and to inform 2010 Census planning. Findings from the Census 2000 TXE Program reports are integrated into topic reports that provide context and background for broader interpretation of results.

Sid Schneider, David Cantor, Paul Segel,
Carlos Arieira, Luu Nguyen
Westat

Larry Malakoff, Project Manager
Statistical Research Division

U S C E N S U S B U R E A U
Helping You Make Informed Decisions

CONTENTS

CONTENTS (continued)

List of Appendixes

List of Tables

CONTENTS (continued)

List of Tables (continued)

List of Figures

CONTENTS (continued)

List of Figures (continued)

CONTENTS (continued)

List of Figures (continued)

EXECUTIVE SUMMARY

The Response Mode and Incentive Experiment investigated the impact of three computer-assisted data collection techniques on the response rate and data quality in Census 2000. Households participating in the study were randomly assigned to six panels and to a control group. The households in the six panels were given the choice of providing their Census 2000 data via the usual paper forms or by an alternate computer-mediated response mode. Half of these panels were offered an incentive, a telephone calling card good for 30 minutes of calls, for using the alternate response mode.

The results suggested several patterns relevant to the use of computer-mediated response modes in the decennial census.

Effects of Mode on Response Rates

The overall response rate increased when the Computer-Assisted Telephone Interview or the Internet was offered as an alternate response mode.

The results for the Interactive Voice Response are more difficult to interpret because a portion of the sample in the Interactive Voice Response no-incentive panel either received the census form late or did not receive it at all. It was possible to identify some, but not all, of the geographic areas affected by this problem. The analysis of these response rates is limited by the suspicion that the rates for this panel are depressed by a problem with the mailing. It is also limited because areas identified as being affected by the problem were eliminated from the analysis, thereby producing results for a subnational sample.

With this caveat, the results indicate that the response rate for the Interactive Voice Response panel was not statistically different from the control (mail only) group, and was lower than the response rates for the Computer-Assisted Telephone Interview and Internet groups.

Effects of the Incentive on Response Rates

The incentive brought about very large increases in the use the alternative computer-assisted response mode. That is, the incentive caused large numbers of respondents to switch from the paper form to an alternative response mode.

In the Computer-Assisted Telephone Interview and Internet conditions, the incentive actually brought about a decrease in the overall response rate; that is, when the incentive was offered, the response rate via the alternative response mode increased, but the response rate via paper forms fell, more than offsetting that increase.

Item Nonresponse

By a wide margin, the highest item nonresponse rates occurred for the Interactive Voice Response mode. A large proportion of the missing data for this mode was attributable to respondents hanging up the telephone before the interview was complete. Other sources of nonresponse were issues related to use of the paper form, the ethnicity question and the race questions.

The paper forms method had lower item nonresponse rates, followed by the Internet. The Computer Assisted Telephone Interview had a relatively low item nonresponse rates, probably because the interviewers were trained to avoid missing data.

The data could not suggest why the item nonresponse rate was lower for the Internet data collection mode than for the paper forms mode. Respondents were permitted to omit their responses to most questions in either of these modes.

Satisfaction with the Interactive Voice Response Mode

Respondents whose households had a large number of people tended to be dissatisfied with the Interactive Voice Response survey. They tended to find it confusing, disliked the amount of time they were given to answer the questions, and required a relatively long time to provide their data. The length of time needed to fill out the questionnaire, which was highly correlated with the number of persons in the household, was also associated with dissatisfaction.

Conclusions

- The Computer-Assisted Telephone Interview brought about a small but statistically significant improvement in the overall response rate. It also had a low item nonresponse rate. However, in the context of this experiment, it entailed substantial cost for hardware, software, and programmer and interviewer time.

- The Internet is an attractive alternative data collection mode for the decennial census. The Internet mode yielded relatively high data quality. The primary additional cost associated with the Internet survey involves the development and maintenance of the software and hardware. The benefits of this data collection method may outweigh these costs.

- The implications of this study are complex for the use of the Interactive Voice Response mode in the decennial census. Data quality was the lowest for this data collection mode. Respondents appeared to dislike lengthy surveys with this method. Nonetheless, this mode would seem to be an appealing way to reach persons with limited literacy skills. The costs involved include the hardware, programming, speech recognition software, and telephone expenses.

- The calling card was very effective in promoting the use of the alternative response mode. The incentive tended to redirect households that would have responded by mail anyway to the alternate computer-mediated response mode. This effect, however, may be partly attributable to the ability of the colorful inserts to call attention to the calling card. Research is needed into the best ways to present the response mode alternatives through either the letter or an insert.

- The impact of the calling card may not justify its cost. In the Internet and Computer-Assisted Telephone Interview conditions, the incentive may have brought about an increase in responding via the alternate mode, but this increase was offset by decreases in responding by mail.

1. BACKGROUND

New computer technologies are transforming the field of survey data collection in many ways. The expansion of the Internet has made web-based surveys more feasible. Improved software has made computer-assisted telephone interviews (CATIs) easier to implement. New speech processing systems have made large-scale Interactive Voice Response (IVR) surveys possible. Each of these computer-assisted data collection methodologies may assist the U.S. Census Bureau in fulfilling its mission. However, the potential benefits of Internet-based surveys, CATI, and IVR surveys can be realized only if large numbers of respondents are willing to answer survey questions using these computer-assisted data collection methods. The objective of the Response Mode and Incentive Experiment (RMIE) was to investigate the potential role of these technologies in Census 2000.

The specific goals of the present study were:

- To assess the public's willingness to provide census data using these computer-mediated data collection methods;
- To evaluate the quality of the data collected using these methods; and
- To study the ability of incentives, in the form of telephone calling cards to promote the use of these computer-mediated methods.

1.1 Summary of the RMIE Design

The RMIE had three basic components. The first was the initial mailout. Census 2000 forms were delivered to all households in the United States beginning on or about Monday, March 13, 2000. Five-sixths of these households received the short version of the form. A sample of the households that received the short form had been randomly selected, prior to the mailout, for the RMIE. This sample was stratified according to each household's geographical location into two areas. The low coverage area was comprised of sections of the United States with high concentrations of non-White residents and renters, two groups that are associated with low response rates. The high coverage area comprised the rest of the nation.

Some of the households in the random sample served as the Census Control Group (CCG); each of these households received a form and letter identical to those used in the national Census 2000 mailing. The rest of the households in the sample did receive special instructions, giving them the choice of providing their census data either by filling out the paper form, or by using a computer-assisted method:

- One subsample of the households was given the option of providing their census data via a CATI.

- A second subsample was given the option of providing their census data via an IVR system. The IVR census questionnaire is called the Automated Spoken Questionnaire (ASQ).

- A third subsample was given the choice of providing their data on a web-based survey.

Half of the households in each of these three experimental conditions were offered telephone calling cards as an incentive to use the computer-assisted method to report their census data. After households provided their census data using the computer-assisted method, their calling cards were given value worth 30 minutes of domestic telephone calls.

The second component of the RMIE was an operation to follow up with the nonrespondents of the CCG. Households in the CCG that failed to mail back their census forms—that is, the nonrespondents to the initial mailout—were given the opportunity in late April to provide their census data using one of the three computer-assisted methods. Half of these nonrespondents were offered a calling card as an incentive to use a computer-assisted method. Thus, the design of this nonresponse (NR) phase of the RMIE was very similar to the design of the initial mailout component. The results of the NR phase of RMIE were reported by Guarino (2001).

The third component of the RMIE was an Internet Usage survey. This telephone survey involved a sample of the households that were offered the opportunity to fill out the Internet version of the census short form in the initial mailout in March 2000 but either mailed in their data on the paper form or called the operator assistance (OA) number and provided their census data to a telephone interviewer. The Internet usage survey explored the reasons why these households chose not to provide their information using the web-based survey (this component of RMIE is covered in Bouffard & Guarino, 2001).

These three components are shown in Figure 1.

This report describes the results of the first component of the RMIE, involving the initial mailout of the census forms in March 2000. The methods section of this report explains the experimental design, the CATI, ASQ, and Internet data collection methods, the OA toll-free number, the calling card incentives, and the data analysis procedures. The results section of this report presents the observed response rates—that is, the proportions of households that provided their census data via a computer-assisted method and on the paper forms. The results section also discusses data collected about the usability and the respondents' opinions of the ASQ method.

Components of the response mode and incentive experiment (Figure 1)

I. Initial mailout component

No incentive panels: Alternate modes
- Panel 1: Computer-Assisted Telephone Interview (CATI)
- Panel 2: Automated Spoken Questionnaire (ASQ)
- Panel 3: Internet

Incentive panels : Alternate modes
- Panel 4: CATI
- Panel 5: ASQ
- Panel 6: Internet

Census Control Group (CCG): No incentive, no alternative mode

II. Nonresponse Follow Up (NR Phase)

Households in the census control group who failed to provide their census data in the initial mailout were included in a Nonresponse Phase study with six panels, similar in design to the initial mailout study.

III. Internet Usage Survey

Households in Panel 3 and Panel 6 who provided their census data, but not via the Internet questionnaire, were surveyed to learn their reasons for not using the Internet.

1.2 Web-Based Surveys

In web-based surveys, respondents access a survey web site, type in codes to authenticate their identities, and fill out the survey. They respond to multiple choice questions by clicking on their responses. They respond to text-entry questions by typing their responses in a field.

Web-based surveys have several advantages over paper-and-pencil and face-to-face interviews. Potentially, the survey software can automatically prevent respondents from making common errors, such as selecting more than one response in "choose one" multiple choice questions, or submitting surveys with contradictory or missing information. The software can also prevent respondents from filling out the survey more than once. In addition, the software can automatically follow the branching logic, so that the respondent receives only the appropriate questions. However, computer-assisted data collection methods probably cannot reduce the rate of "don't know" responses or refusals (Baker, Bradburn, and Johnson, 1995).

Web-based surveys are often convenient for both respondents and researchers. The respondents can submit their information at any time, from any computer with Internet access. Since no interviewer is involved, web-based surveys ensure that all respondents receive precisely the same

information and instructions for the survey. The data are available to the researchers immediately after the respondents fill out their surveys. No transcription is needed, eliminating delay and potential errors.

However, web-based surveys also have several drawbacks. Most importantly, most American households still do not have Internet access, even though the proportion of households with Internet access rose from 26.2 percent in December 1998 to 41.5 percent in August 2000. In December 1998, 15.8 percent of Americans (individuals, not households) used the Internet at home only, 10.5 percent at work only, 6.5 percent both places, and 67.3 percent not at all. In August 2000, 25.0 percent of American individuals used the Internet at home only, 8.7 percent at work only, 10.7 percent both places, and 55.6 percent not at all (National Telecommunications and Information Administration, 2000). Those who do not have access to the Internet at home or work may have access at a library, community center, or "Internet café," but these locations can be inconvenient. Some elderly, infirm, and physically challenged persons who do not have access to the Internet at home may have particular difficulty gaining access to the Internet outside of their homes. In addition, some persons may not feel confident that they have the skills needed to fill out a web-based survey.

1.3 CATI Surveys

With inbound CATI, the respondents call a toll-free number to reach an interviewer. The CATI system assists the interviewer in collecting the data. The software automatically follows the branching logic. It displays the wording required for each question on the interviewers' computer screens, along with the codes for the acceptable answers. The interviewers read the questions and enter the responses. The interviewers can document any unusual responses or situations, such as broken-off interviews. In this way, interviewers administer the survey and record the data in a standardized manner with a minimum of variation among interviewers. The data are available quickly with no need for transcription. The CATI system also provides interviewers with an automated method to keep track of each respondent's interview status. For example, in an inbound CATI study, the interviewers can readily identify the respondents who have called and the respondents who are scheduled for followup calls.

According to 1990 Census data, 94.8 of American households had a telephone that year. The rate was lower in some regions; for example, only 87.4 percent of Mississippi households had a telephone. Some of the households without telephones would have difficulty providing data using inbound CATI. However, a large proportion of these households may have access to a neighbor's telephone, pay telephone, or business telephone.

1.4 ASQ

An IVR survey enables respondents to take a survey on the telephone by interacting with a "talking computer." The respondents hear digitized voice files containing the survey instructions and questions. The respondents answer by speaking. The computer uses speech recognition software to transform the respondents' spoken replies into text that can be stored in a database. Contemporary speech recognition software is speaker-independent; that is, it attempts to recognize spoken words regardless of the accent, gender, or identity of the speaker. To this extent,

IVR questionnaires emulate interviewer-administered surveys. For example, with an ASQ, the respondent might hear, "What is your sex?" The computer would accept the respondent's spoken replies, such as "male" or "I am a woman" or "I'm female." Of course, the IVR system can also accept responses entered with the familiar "touch tone" buttons. For example, the respondent might hear, "Please enter your telephone number, using the touch-tone keys." The respondent would reply by pressing the buttons.

IVR surveys offer survey researchers several advantages. Like CATI or web-based systems, an IVR system automatically follows the survey branching logic, and stores responses in a database, eliminating transcription and its associated costs and errors. Researchers may also elect to have some responses transcribed if they are concerned that the voice recognition software may introduce errors of its own.

Much like a web-based survey, an IVR survey allows respondents to provide data at any time. However, unlike a web survey, an IVR survey does not require respondents to have good literacy skills or to be comfortable with the Internet. An IVR survey may be a relatively economical data collection method because it requires many fewer operators than a CATI survey. As IVR systems are more widely used, the extent of such cost savings will become clearer.

IVR applications are quickly becoming more numerous and popular. Callers are likely to encounter IVR automated-attendant systems when they telephone the customer service departments of airlines or financial service companies or other firms that handle a large number of customer inquiries. IVR applications may be perceived as interesting or "high tech" data collection methods. Conceivably, some respondents prefer this data collection method over others and may be more willing to use it when given the opportunity.

However, IVR surveys also have drawbacks. Speech recognition technology is new and has not yet been perfected. These systems can be error-prone, particularly when they are used for surveys that contain questions for which there are many possible responses. Moreover, most people have little experience using these systems and may not feel completely comfortable speaking to an automated device.

1.5 Rationale and Previous Research

Respondents who are unable or unwilling to use the paper form to provide their census data might use alternative modes of responding, if those alternatives were made available. Alternatives to the paper form might therefore bring about an increase in the response rate and improved data quality for the decennial census. Respondents who fail to fill out the paper forms might provide their data on the Internet or by telephone to an operator or to an IVR application, if the respondents were given the opportunity.

Respondents with limited literacy skills might be especially likely to prefer to provide their data by telephone. Results of the National Adult Literacy Survey has suggested that the reading proficiency of about 30 million English-speaking Americans is at the lowest defined level in that survey. The reasons for their limited literacy skill are varied. About one-fifth of this group have visual difficulties, about one-quarter have cognitive or physical impairments, and most of the remainder have minimal educational attainment (Kirsch, Jungeblut, Jenkins, and Kolstad, 1993).

Many people who have limited literacy skills have particular problems filling out forms. Telephone-based data collection methods could allow these individuals to provide their census data without having to read or fill out any forms. With the ASQ and inbound CATI available, respondents need only be able to provide their Census 2000 data verbally in response to a series of spoken questions.

The results of the Survey of Census Participation, which studied the 1990 census, suggest that some people failed to provide their census data that year because they believed the forms required a great deal of time or because they feared that the Census Bureau might fail to keep the information confidential (Couper, Singer, and Kulka, 1998; Singer, Mathiowetz, and Couper, 1993). The Census Bureau takes measures to assure the public that census data are kept strictly confidential. However, sometimes assurances of confidentiality can backfire by arousing respondents' suspicions and actually depressing the response rate (Singer, Hippler and Schwarz, 1992).

Some respondents may regard computer-assisted data collection techniques as less burdensome and more confidential than paper forms. In one study, respondents rated an IVR application as the "easiest" way to provide personal data (Turner, Miller, Smith, Cooley, and Rogers, 1996). Several studies have suggested that respondents are more willing to reveal personal data with a computer-mediated data collection system than with human interviewers or with paper forms (Turner, Ku, Rogers, Lindberg, Plaeck, and Sonenstein, 1998). Perhaps the ASQ might be acceptable for some of the people who avoid paper forms.

1.6 Rationale for the ASQ Usability and Satisfaction Study

The ASQ used in this study included several questions at the end that were intended to measure the respondents' satisfaction with the interview. The responses to these questions could demonstrate the acceptability of the data collection procedure to the respondents. In addition, the computer system logged the length of time that the respondents required for the interview, and the number of times that the system had to repeat questions in order to obtain an unambiguous response.

Speech recognition technology has not yet been widely used as a computer-assisted survey data collection tool. Developers have instead concentrated on applications with larger markets such as automated call routing systems, voice interfaces for information services, and voice interactive devices in automobiles. Nonetheless, the results of the 1995 census test, a preparatory exercise for Census 2000, suggested that respondents may be willing to use an IVR application and that IVR systems may be able to collect data accurately. That 1995 test further suggested that the manner in which the IVR application was designed was crucial to its acceptance.

Ron Cole and his colleagues at Oregon Graduate Institute (OGI; Cole, et al., 1995) designed an IVR application to collect census short form data for the 1995 Census test. They called this system an Automated Spoken Questionnaire or ASQ. The OGI researchers found that the speech recognition software they used was more than 98 percent accurate in recording the responses to questions that required simple answers like "yes" or "no," and "male" or "female." However, for more complex questions, such as those that required a date as a response, the accuracy fell to less

than 75 percent. Those levels of accuracy were comparable to those reported by researchers who studied other IVR survey applications (Blyth, 1997).

Cole and his colleagues found that respondents generally preferred IVR data collection over paper-and-pencil forms. However, several factors influenced respondent satisfaction. For example, Cole's research team initially designed their system so that when the speech recognition software could not understand a response with a high level of confidence, the system asked the respondents to verify their responses. This procedure was time consuming and did not resemble typical person-to-person conversation. As a result, respondents tended to be dissatisfied with those systems. The OGI researchers changed their system so that when the speech recognition software could not understand a response with a high level of confidence, the system simply asked the question again. If the speech recognition software still could not understand the response, it flagged the speech file for transcription later. Respondents were much more satisfied with this more time-efficient, natural-sounding procedure.

Cole and his colleagues also found respondents preferred systems that presented instructions quickly and that did not require the respondents to perform time-consuming tasks such as entering their 22-digit census identification number. Respondents from large households also appeared to be relatively dissatisfied with the ASQ system, perhaps because of the time they needed to enter data for their many household members.

Before speech recognition technology can be deployed in the decennial census, more needs to be known about the factors that influence its acceptance by the public. This study therefore included measures of the usability of the IVR system and the respondents' satisfaction.

2. METHODOLOGY

A total of 35,376 households were randomly selected for this study from the Decennial Master Address File (DMAF) developed for Census 2000. All of these households were from the 94.3 million households in mailout/mailback areas. Households in list/enumerate and update/leave areas (U.S. Census Bureau, 1999) were not selected. All were scheduled to receive the short form.

Because of the schedule for this study, these households were selected before the U.S. Census Bureau conducted several operations designed to improve the coverage of the DMAF, such as the U.S. Postal Service check. Households that were selected for the Accuracy and Coverage Evaluation initial and final samples were excluded from this study.

2.1 Experimental Design

Of the households selected for this study, 15,737 were randomly dispersed among six panels in a three by two, fully factorial design. The first factor, response mode, had three levels: CATI, ASQ, and Internet. The households were given the choice of providing their census data either via U.S. mail on the usual paper forms, or via their assigned computer-assisted response mode.

The second factor, incentive, had two levels: incentive and no incentive. Households in the incentive condition were rewarded for using a computer-assisted response mode to provide their census data, while those in the no-incentive condition were not. The reward was a telephone calling card.

The six panels and the number of households assigned to each were as follows:

Panel 1: CATI with no incentive	2,621
Panel 2: ASQ with no incentive	2,621
Panel 3: Internet with no incentive	2,626
Panel 4: CATI with incentive	2,621
Panel 5: ASQ with incentive	2,625
Panel 6: Internet with incentive	2,623

Mailings. The Census Bureau mailed a short form for Census 2000 and a cover letter to each household in this study around March 13, 2000, at the same time that census forms were mailed to all households in the nation. Appendix A contains copies of the RMIE mailings. No household in this study received the census long form. The cover letter explained that the household could provide census data in either of two ways. First, the household could mail in the data in the usual way, using the paper form. Alternatively, the household could use a computer-assisted method. The cover letters to panels 1 and 4 explained that the household could provide data over the telephone by dialing the toll-free number (877) 53-COUNT. The letters did not mention that the data would be collected by a CATI operator. The cover letters to panels 2 and 5 explained that the household could provide data by telephone by calling the toll-free number (877) COUNT-10. The letter did not mention the data would be collected by an IVR system. The cover letters to panels 3 and 6 explained that the household could provide data via a web-based questionnaire available at ***www.2000.census.gov.***

The mailings to panels 4, 5, and 6 (the incentive panels) contained an insert, printed in color on heavy stock paper. A calling card was attached to the insert with weak glue (known as "tipping glue") so it could be easily peeled off. The cover letter and insert explained that if the household provided its census data using the computer-assisted method, the calling card would be activated, giving it value worth 30 minutes of domestic calls.

The paper census forms sent to the households in all six panels provided the toll-free number (877) 8-CENSUS for any questions. This number was different from the toll-free help line number that appeared on standard Census 2000 forms received by households that were not in RMIE. This source of help and information was called "Operator Assistance" or simply "OA." Operators were available at that number to answer questions both about this study and about Census 2000 generally.

Census Control Group. The remaining 19,639 households that were selected for this study comprised the Census Control Group (CCG). The CCG received mailings that contained a cover letter and a census short form. The mailings did not offer the CCG households the opportunity to provide census data using a computer-assisted response mode, nor did the mailings offer any calling card or other incentive. The CCG served as a group against which the six panels in this study could be compared. In addition, households in the CCG that failed to provide their census data were involved in a Non-Response (NR) phase (see Guarino, 2001).

Because the households in the CCG that failed to provide their census data were included in a NR phase, their census forms listed the special OA number for RMIE, not the standard Telephone Questionnaire Assistance number printed on the Census 2000 short forms that went to households not selected for RMIE. Except for the OA telephone number, the mailings received by the CCG were identical to the official Census 2000 short form and cover letter. When CCG households had questions about the NR phase and called the RMIE OA number, they reached an operator who was knowledgeable both about RMIE and about Census 2000 generally.

Stratification. Each household selected for this study was classified as being from one of two strata: a low coverage area (LCA) or high coverage area (HCA). The LCA was comprised of census tracts with high concentrations of non-White residents and renters, two groups associated with high nonresponse rates. About 19.3 percent of the households in the DMAF in mailout/mailback areas are in the LCA; the HCA consists of the remaining households. In RMIE, households were proportionately selected from the two strata; just under one-fifth of the households in each panel and in the CCG were in the LCA stratum.

2.2 ASQ

Only households assigned to panels 2 and 5 were informed of the ASQ telephone number. Therefore, all calls to the ASQ were only from households in those two panels. The protocol for the ASQ is included as Appendix B.

The ASQ closely followed the paper Census 2000 short form. However, unlike the paper census short form, the ASQ allowed the collection of information about all members of a household, no matter how many there were. The paper short form asked for information about only six persons in the household; it collected only the first and last names of the seventh through the twelfth persons, and no information at all for any persons beyond the twelfth.

The respondent answered all questions on the ASQ by speaking, with the exception of the questions asking for the following data:

- The household's telephone number;
- The household's 22-digit census identification number (from the address label on the mailing),
- The ten-digit calling card number (for panel 5 [the ASQ-incentive panel] only)

The respondents provided this information by pressing the touch-tone keys on their telephones. However, respondents who were not using a telephone with touch-tone keys provided this information verbally.

Computer hardware exists to accept data input entered with a pulse telephone, which generates pulses rather than tones. Pulse telephones can be rotary dial or push button. Consultants to the RMIE from Unisys strongly advised against this hardware because it does not function reliably.

In particular, it tends to accept noise on the telephone line as valid data input, corrupting the data actually entered by the respondents. For this reason, the ASQ system used in RMIE was configured to ignore data that a respondent might attempt to enter with a pulse telephone.

Immediately after respondents entered their 22-digit census identification numbers, the system determined whether the respondents had called the system previously. If a respondent had called earlier, the system transferred the call to a CATI operator who collected any updated information from the respondent. The ASQ system also transferred a call to a CATI operator if the respondent did any of the following:

- Failed to provide the 22-digit census ID when asked;
- Attempted to enter the census ID with a pulse telephone; or
- Entered a census ID that was not in the databases for panels 2 or 5.

The CATI operator would help the caller find the correct 22-digit number and then collect the caller's census data.

Many of the questions in the ASQ required the respondent to reply simply by saying "yes" or "no." The system immediately recognized the respondent's reply. As specified in the protocol for certain questions, the system branched to the succeeding question depending upon the reply.

When speech recognition software attempted to recognize an utterance, it returned a confidence level associated with the recognition attempt. The level was expressed as a percentage, generally between 50 and 100. Recognition attempts with high confidence levels were more certain than attempts with low confidence levels.

If the software returned a confidence level under 70 percent in an attempt to recognize a "yes" or a "no" response, the system repeated the question (as shown in the ASQ protocol in Appendix B). If the software still could not adequately recognize the response in this second attempt, the system transferred the call to a CATI operator, who administered the census questionnaire. Certain other questions, shown in the ASQ protocol, also required responses; if the respondent was silent or the software could not recognize the response above a set confidence level when the question was asked and repeated, the call was transferred to a CATI operator.

When calls were transferred this way, the CATI operator started from the beginning of the questionnaire, even if the respondents had already entered some data using the ASQ system. However, the ASQ system did provide the CATI operators with the respondents' 22-digit census identification numbers, if the respondents had entered them. The respondents did not have to provide this number again. If no CATI operator was available at the time that the call was transferred, or if the transfer occurred after CATI working hours, the respondent heard a recorded message, left a name and telephone number, and received a call from a CATI operator later.

Some questions on the ASQ, such as "Please tell us the month, day and year this person was born" required spoken responses that were more complex than a simple "yes" or "no." The system was not programmed to recognize these responses immediately after they were given. Instead, the system recorded these responses so they could be transcribed soon afterward. The

CATI operators transcribed these recorded responses during periods when no CATI calls were arriving.

At the end of the ASQ, the respondents were given the opportunity to change any of their responses to any question. The transcriptionists listened to these changes and altered the data accordingly.

The system could detect when a respondent hung up the telephone before the interview was complete. Therefore, the system did not log "no response" to the questions after the respondent hung up; it simply stopped recording data.

The ASQ was available 24 hours a day. A more technical description of the IVR system that administered the ASQ is available in Appendix C.

2.3 CATI

Persons from households that were selected for this study could reach a CATI operator in three ways:

- Calls to the ASQ system were transferred to a CATI operator when the speech recognizer could not adequately recognize the respondent's responses to certain questions, or when the respondent entered a census identification number that was invalid or that belonged to a household that had already provided data.

- Households in panels 1 and 4 could dial the toll-free number to reach a CATI operator.

- Respondents in households in any panel could call the OA telephone line and offer to provide their data. Even though the OA number was offered primarily to help respondents with questions about this study or about the census generally, some respondents did call the OA number and ask to provide their census data. The OA operator would transfer these calls to a CATI operator who would collect the data, without regard to whether the caller was assigned to a CATI, ASQ, or Internet panel.

Callers would hear a recorded message if they reached CATI during the late night or early morning or when all operators were unavailable. The message asked the callers to leave their names, telephone numbers, and the times that they might be available for a return call. A CATI operator later called the respondent to collect the census data.

At the start of the interviews, the CATI operators first ascertained whether the caller could speak English. If the caller could speak only Spanish, the operator transferred the call to a bilingual operator. If a respondent who spoke neither English nor Spanish called, the CATI operator could not collect any data. Since no communication was possible with these few callers, they were not considered respondents, and had no follow-up contact. If the caller could speak English, the operator began the interview by asking the caller to read the 22-digit census identification number from the mailing label. The operator administered the CATI interview after verifying that the identification number was from a household in this study. The content of the CATI interview closely followed the content of the Census 2000 short form. However, like the ASQ, the CATI

interview collected complete information about all persons in the household, no matter how many persons lived there. The protocol for the CATI interview is included as Appendix D.

If a respondent had to break off a CATI interview midway, the interviewer would schedule a specific time to complete the interview. The CATI interviewer would telephone the respondent at the scheduled time and finish collecting the census data.

The English-speaking CATI operators worked at Westat's Chestertown telephone research facility while the bilingual CATI operators worked at Westat's Rockville telephone research facility. The CATI software was developed in Blaise, a commercial off-the-shelf survey development system. Blaise CATI software allows operators to use their keyboards, rather than mice or other pointing devices, to enter responses and move among the data entry screens.

2.4 Internet Questionnaire

The Census Bureau staff developed and provided the Internet-based questionnaire for the RMIE. Respondents answered multiple-choice questions in the questionnaire by clicking the appropriate radio buttons and checkboxes. They answered text-entry questions by typing their answers into response fields. The questionnaire screens were designed to resemble the short form paper questionnaire. The screens were not programmed with any branching logic or data validity checks. The Internet survey was available 24 hours a day. A printout of the survey appears in Appendix E.

At the start of this study, the *New York Times* and other news outlets, acting independently, informed the public of the web address used to collect data in the RMIE. President Clinton also revealed the web address, during a weekly Saturday radio broadcast. With the release of that information, all households in this study, even those assigned to the CATI or ASQ condition, were able to fill out their census short forms on the Internet. Households that were not in this study could also fill out their census short forms on the Internet.

2.5 OA

OA operators fielded questions both about the census generally and about the RMIE. Sometimes, the OA operators needed to know a caller's panel assignment to answer the caller's question about the RMIE. The operators could determine the panel by asking the callers for their 22-digit identification number. The operators would enter the number in their computers, which then displayed the panel assignment. In addition, the operators could determine a caller's panel assignment if the caller received a calling card and knew its color. The calling cards for panels 4, 5, and 6 were respectively yellow, red, and blue (although they are reproduced in black and white in Appendix A).

Sometimes callers from panels 4, 5, and 6 called with questions about their calling cards. The OA operator could ask these callers for their ten-digit calling card numbers. Using their computers, the operator could then determine the status of the calling card: invalid number, activated, or not activated.

The OA operators had printed reference material that they consulted before answering the callers' questions. For example, if a caller asked, "How can I get a job as an enumerator?" the operator found this question in an index in the reference material, which provided the appropriate answer. When callers posed questions that the OA operators felt unprepared to answer without further research, the operators recorded the callers' names and telephone numbers and called later with the answers.

In Chestertown, Maryland, operators were assigned to handle either OA or CATI calls. The OA operators worked at stations that were alongside the CATI operators' stations. In Rockville, Maryland, each bilingual CATI operator handled both OA and CATI calls from respondents who spoke only Spanish.

Calls to OA were not logged. Therefore, no data are presented concerning the number and the nature of the OA calls received.

Flow chart. Figure 2 contains a chart showing the flow of calls to the CATI, OA, and ASQ telephone lines. The figure allows the interested reader to view the many ways that a call to these three telephone lines might be routed. As the figure shows, the actual route of any call depended on the time of the call and the requirements of the caller.

2.6 Training for CATI and OA Operators

The CATI and OA operators who worked on RMIE received Westat's General Interviewer Training course to familiarize them with Westat's CATI system and procedures. They also received training specifically for RMIE, learning how to use the CATI system to collect data for the Census 2000 short form or how to handle calls on the OA line. All operators working on RMIE were sworn in according to Census Bureau procedures and instructed about the confidentiality of respondents' data.

During CATI training, the operators learned the content and purpose of the questionnaire and the procedures for collecting the data. The CATI training also included role playing, in which the operators worked in pairs, alternately pretending to be interviewers and respondents.

During OA training, the operators learned the nature of the telephone inquiries that they were likely to receive. They learned how to use printed material from the Census Bureau showing the answers to questions about Census 2000 and about the RMIE. The training also included role playing.

2.7 Calling Card Activation

The calling card activation procedures were designed to be as effortless as possible. At the conclusion of a CATI interview, if the respondent's household was assigned to any panel in the incentive condition (that is, panel 4, 5, or 6), the operator obtained the ten-digit identification code on the reverse side of the respondent's calling card. The CATI software allowed the interviewer to check that the calling card number was valid. The interviewer then activated valid cards. If the calling card number was not valid, the operator helped the caller find the number and read it

provided their data after being transferred from an OA operator, even if the

respondents were assigned to the ASQ or Internet condition. When CATI operators entered a calling card number into their computers, the CATI system automatically called a computer program that checked the validity of the number. If the number was invalid, the CATI operator tried to get a better number from the caller. If the number was valid, the program transferred the number to the Intelicard File Transfer Protocol (FTP) web site. Within one business day, a computer system at Intelicard obtained the numbers from the FTP site and activated the calling cards.

When ASQ respondents entered their calling card number, the IVR system called the software routine that checked the validity of the number. Valid numbers were transferred to the Intelicard FTP web site. If the number was invalid, the system prompted the caller to enter it again, as shown in the protocol (Appendix B).

Although the Internet-based questionnaire could accept census short form data from anyone, it was programmed to activate the calling cards of households in panel 6 only. If persons from households in panel 4 or 5 provided their data via the Internet, their calling cards were not activated.

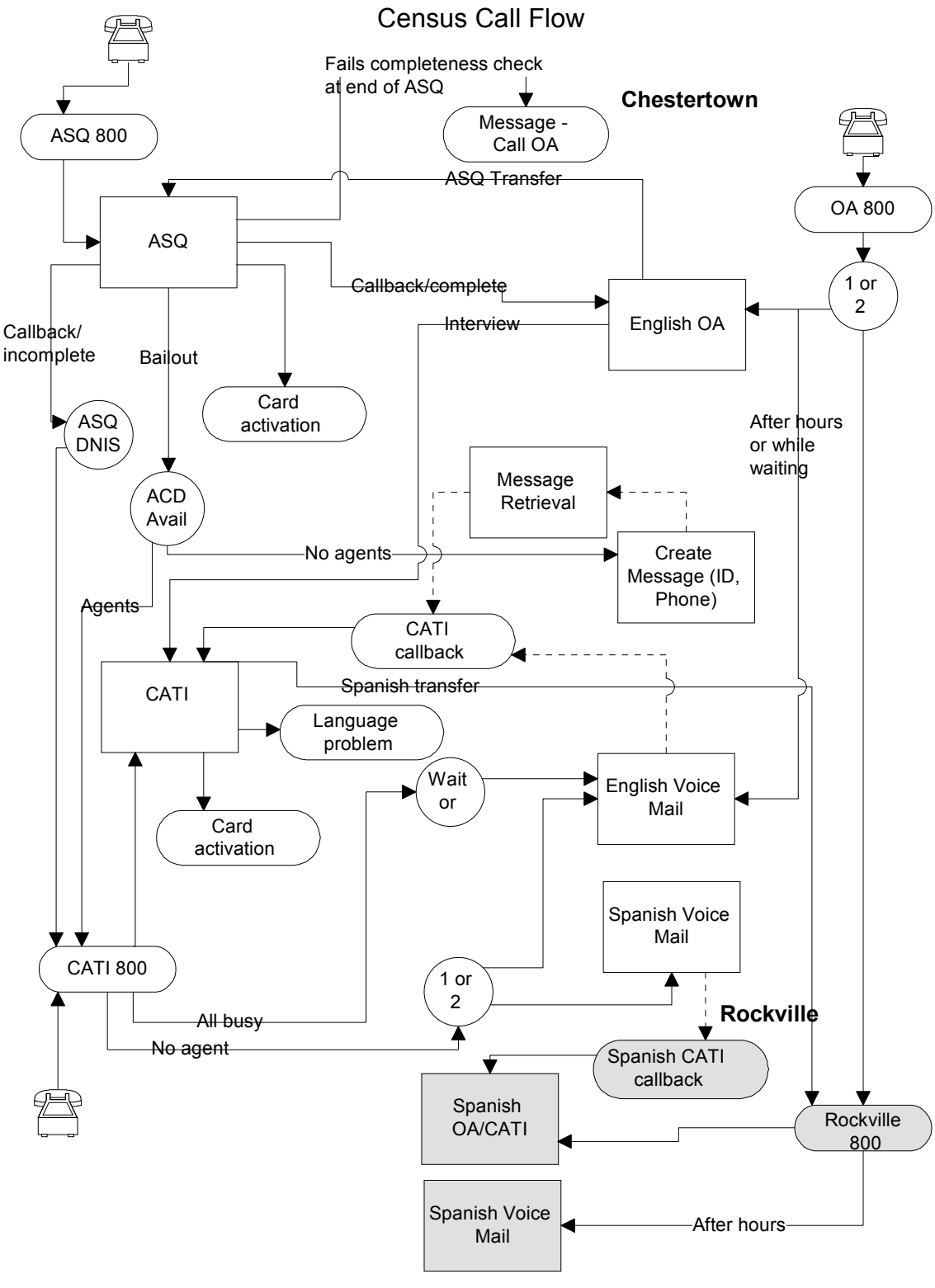
When Internet respondents entered their calling card number, the Internet system generated an email message to Westat containing the calling card number. A computer program (called an email "daemon") automatically opened the email message, extracted the calling card number, and called the computer program that checked the validity of the number. The system transferred valid numbers to the Intelicard FTP web site so that the card could be activated.

On rare occasions, an OA operator received a call from respondents who claimed that their cards should have been activated but still could not be used to make telephone calls. In those cases, Intelicard may have made an error or the caller may have been mistaken. The OA operator had no way to investigate these claims. When this happened, Westat staff manually sent an email to Intelicard to activate the card in question. The number of such emails was not logged.

2.8 Data Reporting

After a CATI operator completed an interview, the respondent's census data were sent to the Census Bureau within one business day. Similarly, after an ASQ was completed, an operator transcribed the voice files that needed transcription, and the data were sent to the Census Bureau within one business day.

The process of transferring Title 13 data to the Census Bureau safeguarded the security and integrity of the data. This process was approved by the Census Bureau and is documented in the


Census Call Flow
Fails completeness check at end of ASQ
Chestertown
ASQ 800
Message - Call OA
ASQ Transfer
OA 800
ASQ
Callback/complete
1 or 2
Interview
English OA
Callback/ incomplete
Bailout
Card activation
ASQ DNIS
After hours or while waiting
Message Retrieval
ACD Avail
No agents
Create Message (ID, Phone)
Agents
CATI callback
CATI
Spanish transfer
Language problem
Wait or
English Voice Mail
Card activation
Spanish Voice Mail
CATI 800
1 or 2
Rockville
All busy
No agent
Spanish CATI callback
Spanish OA/CATI
Rockville 800
Spanish Voice Mail
After hours

RMIE Security Plan and Risk Assessment. First, the data were transferred over Westat's internal network from the CATI and ASQ systems to a system at Westat's Rockville campus that put the data in the format, called "D-1," that the Census Bureau used for all short form data. The system then checked the data to determine whether they were complete enough to be sent to the Census Bureau. To do this, the system applied the completeness criteria used for Census 2000 short form data collection (Bureau of the Census, 2000). The system then created encrypted text files containing the records in D-1 format and placed them on a FTP web site. The encryption was accomplished with Entrust software. Staff from the Census Bureau then accessed the FTP site and transferred the data over the Internet to their own server. They decrypted the files to retrieve the data. Since only Westat and the Census Bureau had the encryption key, this data transfer process was secure.

2.9 Database Management

Westat maintained three databases pertinent to panels 1 through 6. These databases respectively contained the following data:

- Census 2000 data;
- ASQ call data; and
- ASQ response data.

Census 2000 Data. All of the census data collected by the CATI operators and by the IVR system that were sent to the Census Bureau were also stored in this ASCII database.

ASQ Call Data. This database contained the dates and times of every call to the IVR system. It contained a data field that indicated whether or not the call rolled over to the CATI system.

ASQ Response Data. This database contained the respondents' answers to every question in the ASQ, regardless of whether the interview was completed or the call was transferred to a CATI operator or the respondent hung up or was disconnected.

An additional database was obtained from Intelicard. This database showed the extent to which respondents whose calling cards were activated actually used those cards; that is, it contained the value remaining on each activated calling card.

2.10 Quality Control

Several quality control procedures helped ensure the integrity of the census data being collected, stored, and transferred to the Census Bureau:

- CATI operators who encountered problems when collecting data filled out special forms to report the situation. Most of the reported problems were software-related. Supervisorschecked these forms along with the data collected by CATI operators. The

supervisors resolved the problems before the data were prepared for transfer to the Census Bureau.

- Supervisors checked the voice file transcriptions. Any errors were corrected before the ASQ data were prepared for transfer to the Census Bureau.

- At the end of the ASQ, respondents were given the opportunity to record amendments to their earlier answers. (Westat was tardy in making some of these alterations to the Census data, as described in the Limitations section).

- Research staff thoroughly tested the ASQ system before the study began to ensure that the system conformed to the specification provided by the Census Bureau.

- The ASQ system was programmed to check whether respondents had called earlier. Repeat callers were transferred to an OA operator. (As described in the Limitations section, a problem with the database lookup procedure was detected and remedied.)

- Research staff checked the databases, and prepared reports of their contents, each business day throughout the study period. These reports allowed the entire research staff to monitor the progress of the study. Any major difficulties could be quickly identified from these reports.

- Quality assurance procedures were applied throughout the project to the software development, research, clerical activities, data analysis, and report preparation.

2.11 ASQ Usability and Respondent Satisfaction Study

The ASQ concluded with a set of questions to assess the respondents' satisfaction with the data collection method:

- On a scale of one to five, where five means very satisfied, one means very dissatisfied, and three means neither satisfied nor dissatisfied, how satisfied are you overall with the computerized questionnaire?

  If response to the above question was less than three:
  Please tell us what you disliked about the computerized questionnaire after the beep.

  If response to the above question was greater than three:
  Please tell us what you liked about the computerized questionnaire after the beep.

- Were you able to fully understand the computer, yes or no?

  If response to above question was "no":
  Please tell us what you did not understand after the beep.

- Was the computer able to fully understand you, yes or no?

If response to above question was "no":
Please tell us what the computer did not understand after the beep.

- Was there anything about the questionnaire that was confusing or frustrating, yes or no?

  If response to above question was "yes":
  Please tell us what was confusing or frustrating after the beep. . . . . .

- Did you have too much time, too little time, or just the right amount of time to answer the questions?

- Please tell us your suggestions about improving the computerized questionnaire after the beep.

Data Analysis. The usability of the ASQ was measured two ways: (1) by the respondents' answers to the above questions assessing satisfaction with the system, and (2) by measures of the length of time that the respondents required to provide their data. The respondents' answers to the satisfaction questions revealed their opinions of the ASQ. The data analyses included just the questions that called for a numeric response (e.g., a number one to five) or a yes-or-no response. The measures of respondent satisfaction were the following:

- The respondents' ratings of their overall level of satisfaction with the ASQ, as a number from one to five;

- The respondents' ratings of their ability to understand the ASQ, as yes or no;

- The respondents' ratings of the ability of the ASQ system to understand them, as yes or no;

- The respondents' ratings of whether they were confused or frustrated by the ASQ, as yes or no; and

- The respondents' ratings of whether the ASQ afforded them enough time, too much time, or too little time to answer the questions.

The amount of time that the ASQ required was also included as a measure of the usability of the technology as a data collection tool. Two measures were taken of the amount of time required by the ASQ:

- The length of time required by the entire ASQ; and

- The mean time required to answer the survey items. This included the total time required for the system to play the question prompt, the respondent to answer, two seconds to determine if the response is completed (a process called endpointing), and the speech recognition software to compute the response.

A series of regression analyses was planned in which each of the above seven variables served as the dependent variable. In each regression model, the remaining variables were entered as independent variables. In addition, the following variables collected on the census short form also were entered as independent variables:

- Number of people in the household;
- Sex of the respondent;
- Age of the respondent;
- Racial complexity of the household (i.e., the presence in household of at least two people of different races);
- Presence in the household of more than one person who is Hispanic;
- Respondent's race: white, black, or other; and
- Respondent's ethnic origin: Hispanic or non-Hispanic.

The wording of the ASQ was somewhat different from that of the paper form. In the paper form, data were collected first for "one of the people living here who owns, is buying, or rents this house, apartment or mobile home." Data for other people in the household were collected afterward. Thus, the first data to be collected were not necessarily the data for the respondent (i.e., the person who was providing the data). In the ASQ, the first data to be collected were always the data for the respondent. Therefore, the respondent's data could always be quickly identified for these analyses.

Other independent variables in the regression model pertained to the time that the call was received at the call center:

- The hour of the call (Eastern Standard or Daylight Time as appropriate) to the ASQ: morning (6 a.m. to noon), afternoon (noon to 6 p.m.), evening (6 p.m. to midnight), or night (midnight to 6 a.m.); and
- The day of the week of the call to the ASQ.

Additional independent variables in the regression model pertained to the ASQ item response:

- The total number of retries for silence during the call (i.e., the total number of times that the ASQ had to repeat a question because the respondent failed to answer);
- The total number of retries for invalid responses during the call (i.e., the total number of times that the ASQ had to repeat a question because the confidence level reported by the speech recognition software for the respondent's answer was below the threshold set for that question);
- The sum of the total number of retries for silence and the total number of retries for invalid responses; and
- The mean level of confidence for the speech recognizer for all of the respondent's answers throughout the call.

Data included in the analyses. The data for these analyses were collected during every call to the ASQ during both the initial mailout and NR phase, excluding calls that were rolled over (i.e. transferred) to a CATI operator. Thus, the usability and respondent satisfaction study is the only part of this report that involves NR phase data.

3. LIMITATIONS

Several technical problems occurred over the course of this study.

Some difficulties affected the representativeness of the sample in the ASQ panels:

- When the ASQ system began to accept calls, a software problem in a lookup routine caused the system to inaccurately classify all of the callers as ones who had called before. The system therefore failed to administer the ASQ and instead directed the calls inappropriately to the CATI operators immediately after the respondents entered their 22-digit census identification numbers. This problem began with the first call to the ASQ system on Monday, March 13, 2000, and was resolved by 7 a.m. on Thursday, March 16, 2000. The first 115 calls to the ASQ (110 from panel 5 and five from panel 2) were affected.

- A serious problem became apparent when Westat began analyzing the data. The response rate for panel 2 (ASQ - no incentive) appeared to be very low. This inexplicable effect dwarfed all other observed effects and appeared to be an artifact of some error. Moreover, the proportion of mailings returned as Undeliverable As Addressed (UAA) was much lower for panel 2 than for any other panel. Further investigation revealed an apparent problem with the mailout for panel 2. With very few exceptions, no responses were received, nor were any mailings returned UAA, for panel 2 mailings to households in Missouri, Kansas, Nebraska, Louisiana, and Arkansas (the five states whose ZIP Codes start with 630 to 729), Hawaii, Oregon, Washington, and Alaska (the four states whose ZIP Codes start with 967 to 999), and ZIP Codes 39301 to 39648, 60202 to 60490, and 95608 to 95833. No similar problem existed for any other panel. The Census Bureau investigated this situation and found that some responses did in fact come in from households in these areas, but they arrived too late to be counted. Apparently, the mailout to these areas was either delayed or not sent, preventing the affected households from responding before the cutoff time.

  Moreover, a different problem was detected in panels 1 and 3. For those two panels, the proportion of UAA returns was more than two times higher for the state of Indiana than for any other state. In panel 1, Indiana had ten responding households, 6 nonresponding households, and 51 UAAs. In panel 3, Indiana had 13 responding households, no nonresponding households, and 55 UAAs. These UAA rates were by far the highest UAA rates observed for any state in any panel. The UAA rate for the entire nation for panels 1 and 3 were respectively 10.5 and 11.0 percent. For Indiana alone the rates were respectively 76.1 and 80.9 percent.

  Westat examined the data after having removed all data from Indiana, Missouri, Kansas, Nebraska, Louisiana, Arkansas, Hawaii, Oregon, Washington, and Alaska, and ZIP Codes

39301 to 39648, 60202 to 60490, and 95608 to 95833. Chi square analysis revealed that the UAA rate differed among the six panels, even with these areas excluded (chi square = 10.13, df = 5, $p < 0.073$). Further tests revealed that this effect was entirely attributable to panel 2. The UAA rate for panel 2 was significantly lower than the rate for all the other panels (chi square = 7.62, df = 1, $p < 0.006$). No such significant effect was found for any other panel. Thus, even without the ten problematic states and the three problematic ZIP Code areas, the UAA rate for panel 2 was significantly depressed. This finding suggests that problems may still exist with the mailout for panel 2, even after the problematic states and ZIP Code areas are eliminated.

Westat and the Census Bureau discussed these issues and decided that Westat would perform two sets of analyses. One set of analyses involves only four panels; panel 2 would be excluded because of its mailout problems, and panel 5, the other ASQ panel, would also be excluded to maintain a balanced, factorial design. All households in the remaining four panels would be included in this analysis. The problem for Indiana in panels 1 and 3 would be ignored. Inasmuch as the Indiana problem involves UAA rates, not non-response rates, the impact of the problem on the response rates should be relatively minor.

The other set of analyses involves all six panels. However, households from the ten problematic states and the three problematic ZIP Code ranges are excluded from the analyses. The assumption underlying this analysis is that data errors are eliminated by excluding these households. That assumption may not be correct; the depressed UAA rate for panel 2 suggests that problems may still exist even when the ten states and three ZIP Code areas are eliminated. These analyses do not involve a truly national sample, since so much of the country is excluded from the sample. The results should not be generalized to the entire nation.

Some problems affected the completeness of the data collected:

- A problem affected all calls to the ASQ from March 16, 2000, until April 7, 2000. When the speech recognizer found a match to a spoken response but the confidence level was below the threshold level set for that question, the system failed to record the confidence level in the database. However, during this period of time, there were only three calls that contained utterances for which the confidence level was below the threshold. One of those calls had two such utterances, and the other two had one each, for a total of only four utterances with subthreshold confidence levels. Because so few utterances were involved, this problem was ignored in the statistical analyses.

- One respondent apparently put the telephone down or responded to a “call waiting” signal, without hanging up, midway through an ASQ interview. The data for this respondent revealed that the ASQ system repeated questions 98 times because the respondent failed to answer. No other respondent had more than nine retries for silence. This respondent failed to answer any of the questions concerning satisfaction with the ASQ system. The data collected for this respondent were not included in the usability respondent satisfaction study.

Some problems affected the operation of the study, but had no impact on the results:

- For about one week, Intelicard was unable to activate any calling cards. According to the staff at Intelicard, their computers were struck by a virus. When the problem was remedied, Intelicard activated the backlog of calling cards. The OA operators were informed of the situation and assured callers that their cards would be activated when the problem was resolved. Intelicard did not report the number of calls affected.

- Some respondents took advantage of the opportunity at the end of the ASQ to correct the data they had entered during the interview. Westat personnel were tardy in listening to and implementing some of these corrections. Despite this delay, Westat sent all corrected census data to the Census Bureau and made the necessary corrections to all data that were analyzed for this report.

- If an ASQ respondent failed to provide a name for one of the people in the household, the record omitted all the data for one person in the household. Also, some data collected by the ASQ were deleted for calls that rolled over to the CATI system. These problems were corrected easily by retrieving the deleted data from the ASQ response database.

- Early in the study, Westat staff noticed that a large proportion of households in panel 4 (the CATI incentive panel) were providing data using the Internet site, while a large proportion of households in panel 6 (the Internet incentive panel) were providing data using the CATI system. To investigate this situation, OA operators asked callers from panels 4 and 6 about the colors of the calling cards they received. The callers' replies suggested that the mailings to panels 4 and 6 were reversed. Further investigation by the U.S. Census Bureau revealed that someone exchanged the names of the files that contained the 22-digit identification numbers for the households that were assigned to panels 4 and 6 when these files were given to Westat at the start of the study.

  This problem did not undermine the experimental design. The Internet system did collect census data, albeit from the panel that was originally intended to receive the mailing for the CATI-incentive group. The CATI operators also collected data, albeit from the panel that was originally intended to receive the mailing for the Internet-incentive group. Therefore, the analyses could proceed in a straightforward way, simply by reversing the panel assignments. In this report, "panel 4" refers to those households that received the mailing for panel 4, even though those households were originally intended to be panel 6. Similarly, "panel 6" refers to those households that received the mailing for panel 6, even though those households were originally intended to be panel 4.

  However, this problem did have an impact on the respondents who received the mailing for panel 6 who provided their data on the Internet. Because the panels were reversed, the Internet system could not recognize these respondents as being in panel 6 and therefore did not present them with the calling card activation screens where they could enter their ten-digit card number. Their calling cards were not activated. The Census Bureau decided not to manually activate these calling cards. When this problem was discovered, instructions were issued to the Census 2000 Telephone Questionnaire Assistance centers, Intelicard (the calling card vendor), and the Census Public Information Office to give the number for RMIE OA to any RMIE respondents who called to complain. Westat OA operators then could ask the respondents for their calling card numbers and have the numbers sent on to Intelicard to be activated. Similarly, respondents who complained through the official census web site by email received a reply with instructions to call the OA phone number.

The Census Bureau investigated several solutions to activate the calling cards. The first idea was to activate all calling cards sent to households in the Internet-incentive condition. This idea was discarded when it was learned that the only way to do this was to activate all RMIE cards. The cost for universal activation was high and not a viable solution. Senior Census Bureau managers discussed several options to rectify the problem: (1) a special mailing to Internet respondents containing an activated card, and (2) phone calls to the Internet respondents in which we offered to activate their existing cards or mail them an activated card. None of these options were implemented because the few complaints about inactive cards had been resolved and concerns about additional respondent burden. The additional respondent contact posed the risk of the Census Bureau being viewed as invasive and thereby jeopardizing an eventual response during followup operations.

Respondents who provided their data through the CATI system did not experience this problem with the activation of their calling cards. The CATI system was programmed to activate calling cards regardless of whether the respondents were assigned to panel 4, 5, or 6. Therefore, the CATI system activated the calling cards of respondents who were actually intended to be in panel 6 (the Internet incentive panel) who called the CATI system.

4. RESULTS

This section presents the results of the RMIE. Sections 4.1 and 4.2 cover the response rates, Section 4.3 describes the results with respect to item nonresponse, and Section 4.4 presents the results of the usability and respondent satisfaction study.

As discussed in the Limitations section above, the analyses were conducted two ways, because of the apparent mailout problems.

First, in Section 4.1, the results are presented for the analyses involving all households but only four panels. The ASQ panels (panels 2 and 5) are excluded. Part 4.1.1 describes the effects of the incentive on response rates. Part 4.1.2 describes the differences in response rates between the CATI and Internet response mode conditions.

In Section 4.2, the results are presented for the analyses involving the subnational sample and all six panels. Households from ten states and three additional ZIP code ranges are excluded. Part 4.2.1 describes the effects of the incentive on response rates. Part 4.2.2 describes the differences in response rates between the CATI, ASQ, and Internet response mode conditions.

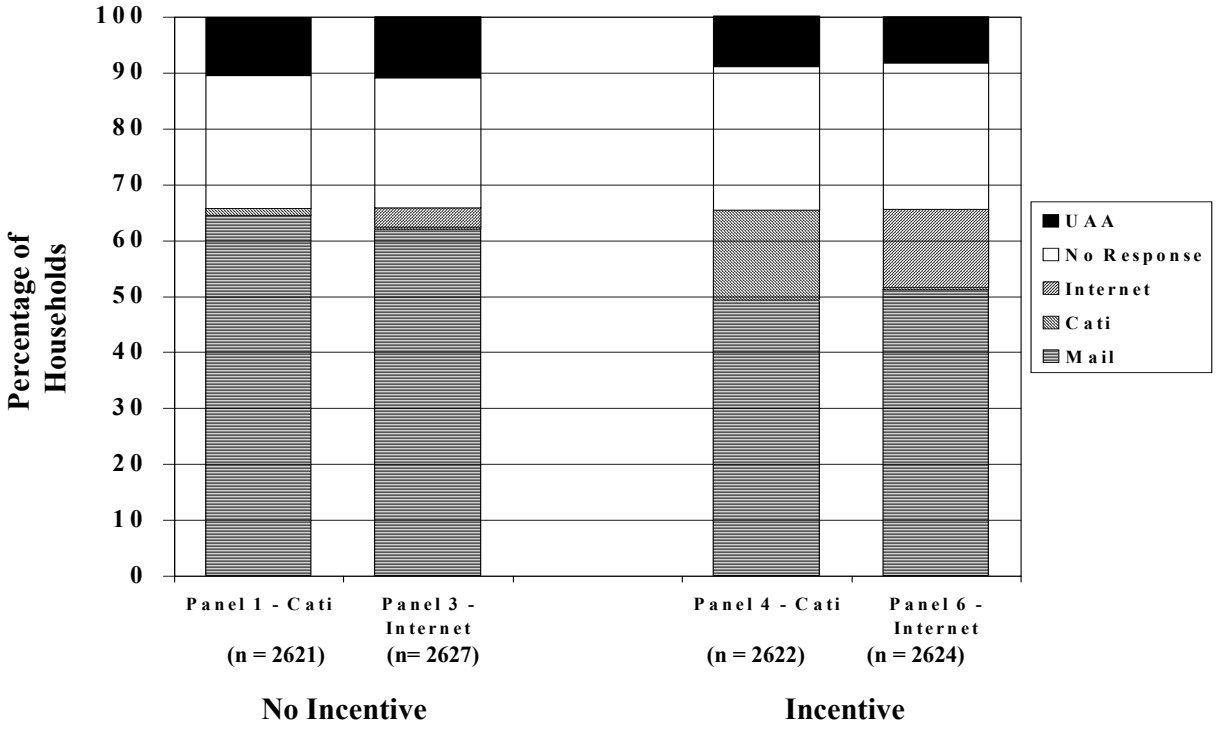
4.1 Effects of the Experimental Conditions on Response Rates: Four Panels, National Sample

These analyses involve only households in panels 1 (CATI-no incentive), 3 (Internet-no incentive), 4 (CATI-incentive), and 6 (Internet-incentive).

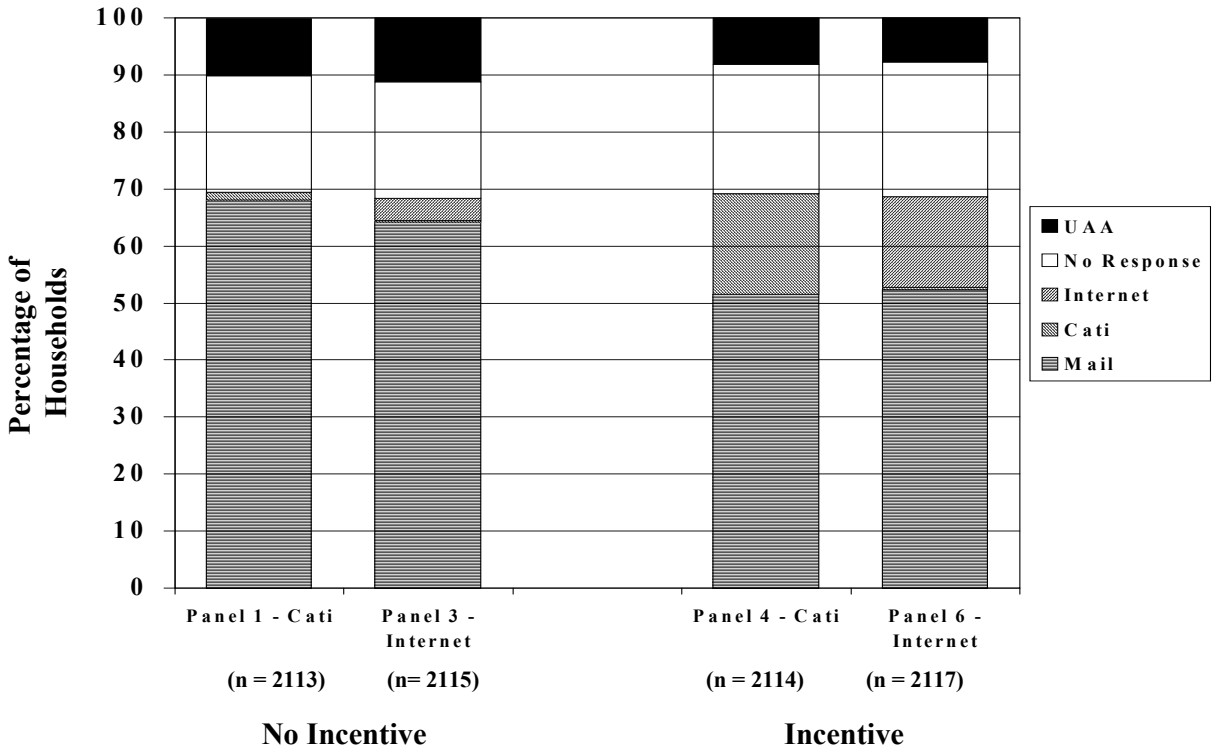
Figure 3 displays the outcome of the mailings to the households in each of the four panels. The figure shows the proportion of households whose mailings were undeliverable as addressed (UAA), and the proportion that did not provide census data, that mailed in their paper forms, and that provided data via CATI or the Internet.

Figures 4 and 5 separate the data shown in Figure 3 by stratum. Figure 4 shows the results for the high coverage area while Figure 5 shows the results for the low coverage area.

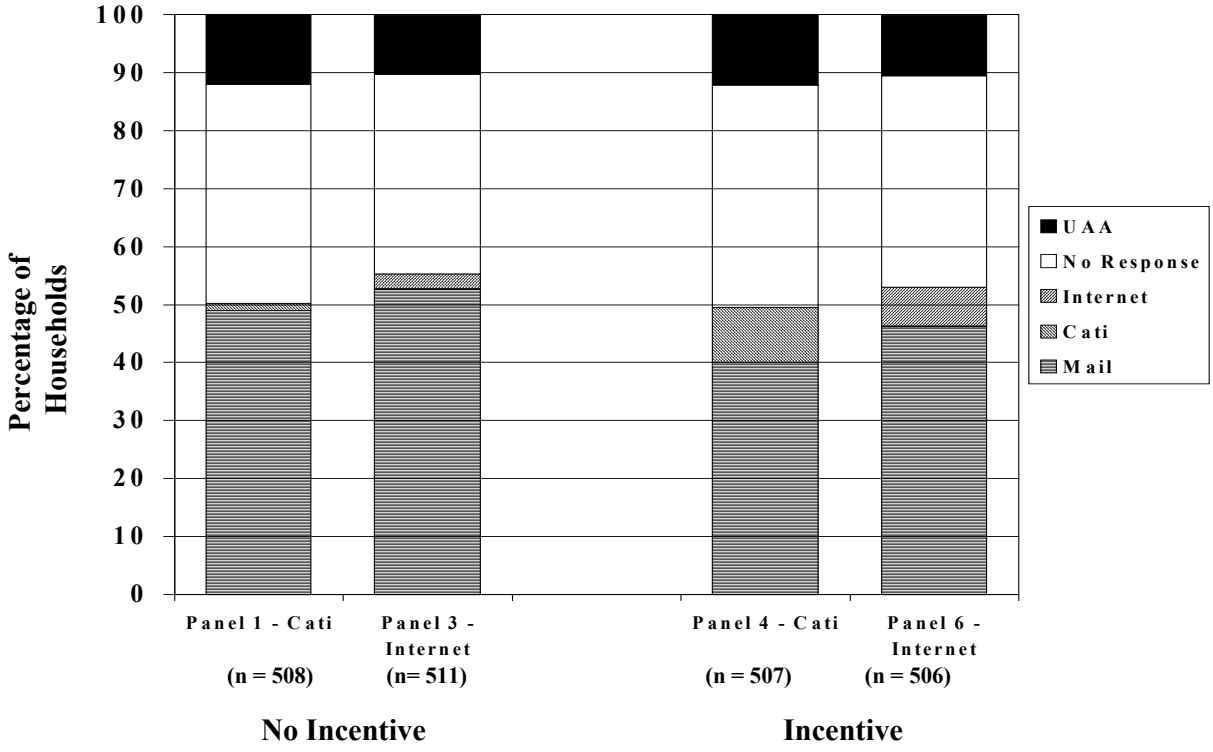
Response to the 2000 Census (Figure 3)



Response to the 2000 Census: High coverage area only (Figure 4)



Response to the 2000 Census: Low coverage area only (Figure 5)



The tables in Appendix G present the data for these figures and all the other figures where the data do not appear on the figures themselves.

The response rates for each panel were computed in two different ways:

- The first computation considered all responses, regardless of the response mode. This computation included responses using the paper form and any responses using the Internet or CATI. The response rates calculated this way were called the "overall response rates."

- The second computation considered only the alternative computer-mediated response modes that were offered in the mailings to the respective panels. Thus, the response rates for panels 1 and 4 considered only CATI responses. Households in panels 1 and 4 were credited with a response when they provided their data to a CATI operator, either by calling the CATI telephone number, or by calling the OA telephone number and being transferred to a CATI operator. They were not credited with a response when they provided their data on the paper forms, or via the Internet. The response rates for panels 3 and 6 considered only responses over the Internet. Households in panels 3 and 6 were not credited with a response when they provided their data on the paper forms or by being transferred to a CATI operator after calling an OA operator. The response rates calculated in this manner were called the "assigned mode response rates."

Figure 6 summarizes these two methods for calculating the response rates. With either method, households were considered nonrespondents if they failed to respond at all, or if they provided data with too many omissions to meet the Census 2000 criteria for a complete response. The Census Bureau furnished Westat with software that identified the responses that did not meet these criteria.

All estimates were weighted by the inverse of the probability of selection. Significance tests were computed using replicate variance estimation methods. Version 4 of WesVar, a Westat-authored computer program, was used to compute the standard error estimates using a jackknife balanced replication procedure. The replicates were formed by sorting the responses by the households' state and county and then designating the 50 systematic samples, each starting with a different integer from 1 to 50, as a set to delete to form a replicate. For example, replicate one was formed by deleting the 1st, 51st, 101st, etc. units and reweighting the remaining 49/50 of the units up to estimate the population total. Replicate two was formed by deleting the 2nd, 52nd, 102nd, etc. units and reweighting the remaining 49/50 of the units.

The response modes counted in computing overall response rates and assigned mode response rates (Figure 6)

	CATI called directly	CATI via transfer from OA	Internet	Mailed paper form
Overall response rate:				
Panels 1 & 4:	✓	✓	✓	✓
Panels 3 & 6:	n.a.	✓	✓	✓
Assigned mode response rate:				
Panels 1 & 4:	✓	✓		
Panels 3 & 6:	n.a.		✓	

Note–n.a. indicates "not applicable."

4.1.1 The effects of the incentive on response rates

This section presents the findings on the effect of the incentive upon the overall response rate and the assigned mode response rate. The section ends with a discussion of these results, addressing these questions:

5. What was the effect of the incentive on the overall and assigned mode response rates?

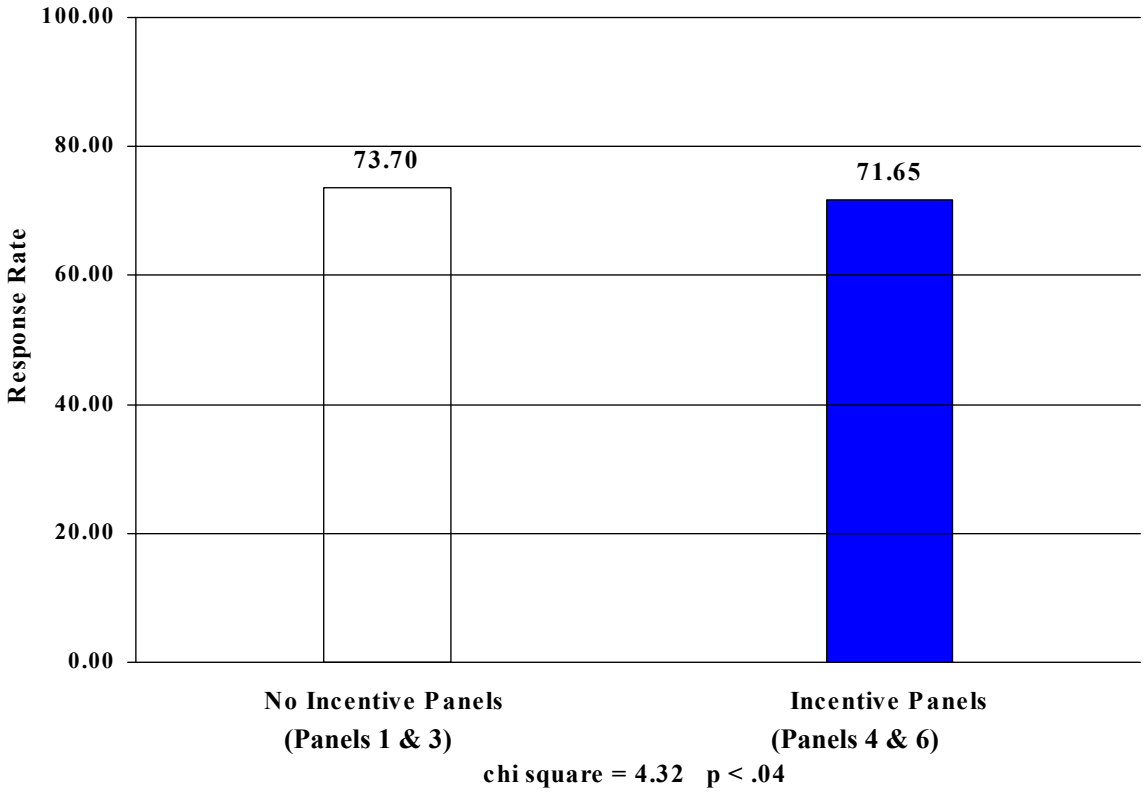
6. Did this effect vary by response mode and by coverage area?

Effects of Incentive on the Overall Response Rate. Figure 7 reveals that the overall response rate was higher for the panels in the no-incentive condition (panels 1 and 3) than for the panels in

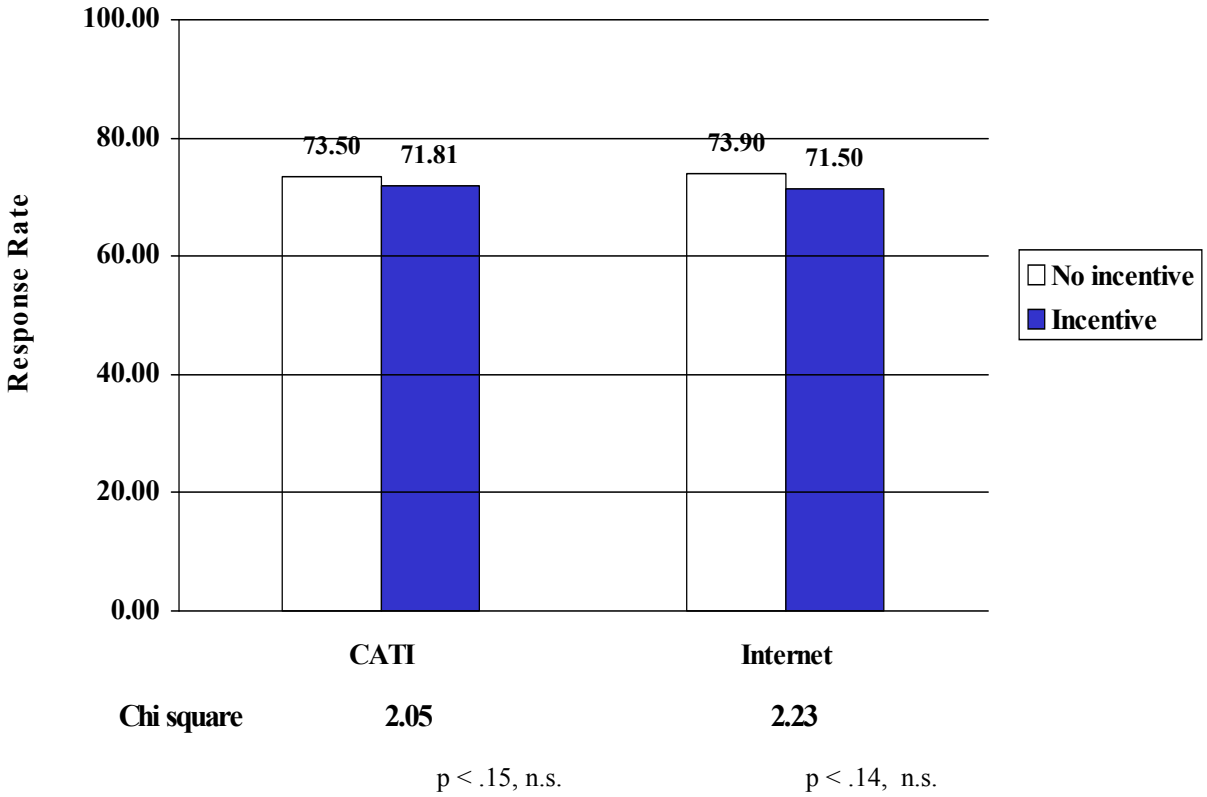
the incentive condition (panels 4 and 6). This difference was small, just 2.05 percent, but it was statistically significant ($p < .05$).

Figure 8 shows that this statistical significance vanished when the CATI and Internet modes were considered separately. That is, the overall response was slightly higher in the CATI no-incentive condition than in the CATI incentive condition, but this effect did not reach statistical significance. Similarly, the overall response rate was slightly higher in the Internet no-incentive condition than in the Internet incentive condition, but the difference escaped statistical significance. In both cases, the difference was less than 2.5 percentage points.

Overall response rate: Combined panels (Figure 7)



Overall response rates by panel (Figure 8)



A logistic regression analysis was carried out to reveal any significant interaction for the overall response rate between the incentive and two other factors—response mode and coverage area. Table 1 displays the results of that logistic regression analysis. The regression model first included all of the main effects and all possible two- and three-way interactions among the response mode, coverage area and incentive conditions; this model was called the "full model." Then the interaction terms that were not statistically significant were removed from the model, one by one, in order of the lowest significance, until all the remaining terms were statistically significant; this model was called the "final model." For this regression model, the reference categories were the Internet (for the response mode factor), high coverage area (for the coverage area factor) and no incentive (for the incentive factor).

Table 1 shows the full model and the final model. The table confirms that the incentive had a significant impact upon the overall response rate; the overall response rate was higher for the no-incentive panels than for the incentive panels. No interaction effects involving the incentive reached statistical significance. The significant CATI-by-low coverage area interaction shown in Table 1 is discussed later, in Section 4.1.2.

Logistic regression analyses of overall response rates (Table 1)

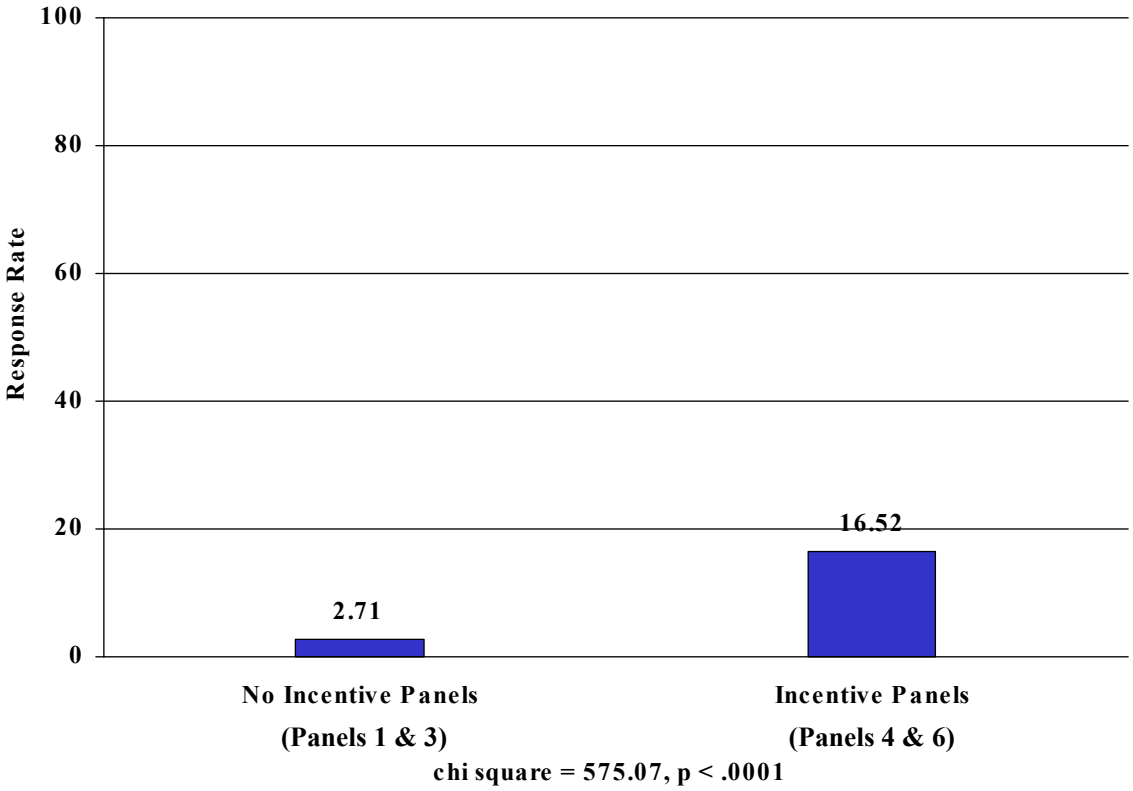
Predictor Variable	Full model Coefficient (Standard Error)		Final model Coefficient (Standard Error)	
Intercept	0.46***	(0.094)	0.47***	(0.066)
CATI	-0.18	(0.119)	-0.15	(0.100)
Incentive	-0.09	(0.152)	-0.11**	(0.051)
High Coverage Area (HCA)	0.74***	(0.105)	0.72***	(0.075)
CATI x Incentive	0.07	(0.170)		
CATI x HCA	0.20*	(0.135)	0.18*	(0.108)
Incentive x HCA	-0.04	(0.151)		
CATI x Incentive x HCA	-0.04	(0.198)		

*p < .10
**p < .05
***p < .001

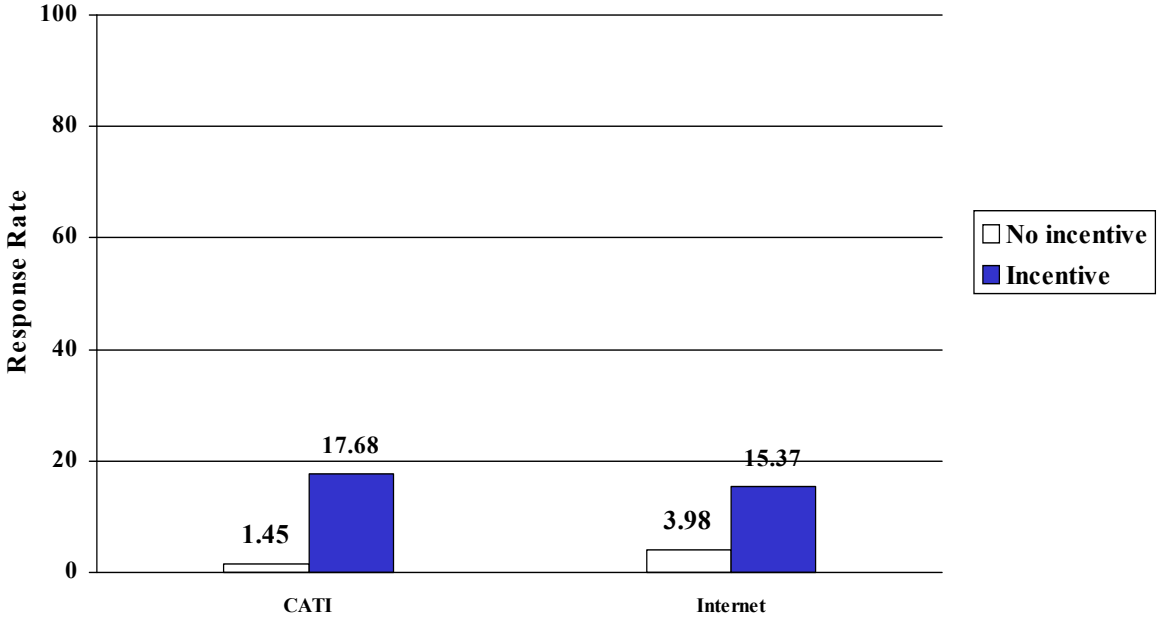
Effects of Incentive on the Assigned Mode Response Rates. Figures 9 and 10 show that the incentive had a large effect on the assigned mode response rate. Figure 9 shows a comparison between the no-incentive panels and the incentive panels. The assigned mode response rate was much higher for the incentive panels. Figure 10 shows that this effect was observed both for panels offered the CATI alternative and for panels offered the Internet alternative. The difference in the assigned mode response rate between the incentive and no-incentive conditions was 16.23 percentage points for the CATI panels and 11.39 percentage points for the Internet panels.

Thus, the incentive increased the likelihood that a household would respond using the assigned alternative response mode, whether that mode was CATI or the Internet. A logistic regression analysis was carried out to test whether this effect for the assigned mode response rates differed by mode or coverage area. This logistic regression was identical in design to the one for the overall response rates, shown in Table 1.

Assigned mode response rate: combined panels (Figure 9)



Assigned mode response rate by panel (Figure 10)



Chi square	**417.94***	**205.26***

***p < .0001**

These results of this logistic regression analysis are shown in Table 2. The assigned mode response rate was lower for the low coverage area than for the high coverage area, as one would expect. There was a statistically significant interaction between the incentive and the mode of response. As Figure 10 suggests, the incentive increased the assigned mode response rate more when households were offered CATI than when they were offered the Internet.

Logistic regression analyses of the assigned mode response rates (Table 2)

Predictor variable	Full model Coefficient (Standard Error)		Final model Coefficient (Standard Error)	
Intercept	-3.54***	(0.305)	-3.80***	(0.141)
CATI	-0.76*	(0.511)	-1.04**	(0.260)
Incentive	1.02***	(0.393)	1.48***	(0.108)
High Coverage Area (HCA)	0.43***	(0.345)	0.72***	(0.120)
CATI x Incentive	1.16***	(0.543)	1.21***	(0.271)
CATI x HCA	-0.33	(0.539)		
Incentive x HCA	0.52*	(0.439)		
CATI x Incentive x HCA	0.07	(0.580)		

*p < .10
**p < .002
***p < .001

Discussion: The Effect of the Incentive on Response Rates. The effect of the incentive can be summarized as follows:

- The incentive offered to the households for responding via an alternative, computer-mediated response mode was associated with a reduced overall response rate, taking into account all modes of response.
- The incentive depressed the overall response rate for both the high and low coverage area households. No significant difference was observed between the high and low coverage areas in this regard.
- The incentive greatly increased the likelihood that the households would choose the alternative response mode.
- The incentive increased the assigned mode response rate via CATI more than it increased the response rate via the Internet.
- The incentive increased the assigned mode response rate for both the high and low coverage area households. No significant difference was observed between the high and low coverage areas in this regard.

These results suggest that the incentive motivated households to use CATI or the Internet to provide their census data. The incentive was slightly (but significantly) more effective in motivating respondents to use the CATI response mode than in motivating respondents to use the Internet. However, this increase in responding by the alternative response mode was offset by a reduction in the response rate by mail; that is, the incentive was associated simultaneously with an increase in responding via the assigned mode, and a decrease in responding by mail. The net result of these two effects was a small reduction in the overall response rates; the overall response rates of the incentive panels were actually lower than the overall response rates of the no-incentive panels. This effect was statistically significant, but not large: 1.69 and 2.40 percent respectively for the CATI and Internet conditions (see Figure 8).

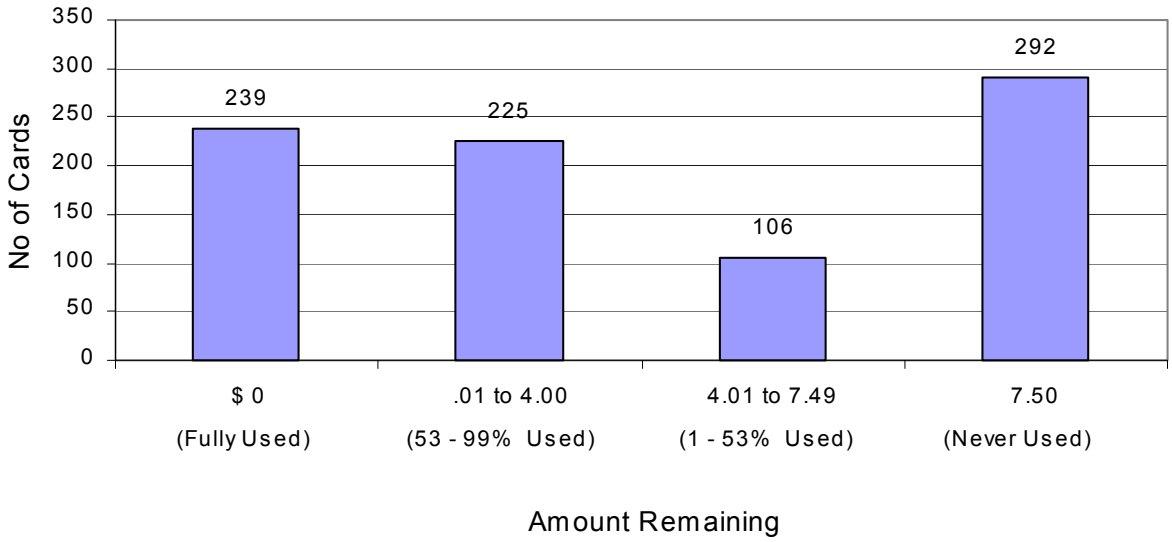
Several explanations may account for these effects of the incentive, simultaneously increasing the assigned mode response rate and decreasing the overall response rate. One speculation is that some respondents attempted to respond by CATI or the Internet when they were offered the incentive but did not succeed in their first attempt. They may have called the CATI during off-hours when the office was closed, or they may have tried to access the Internet while their service provider's servers were busy. Some of these respondents may have then never responded at all because they lost interest or forgot to try again. A second possibility is that some respondents may not have understood the insert and cover letter that explained that they could use the calling card only after they provided their census data. These individuals may have unsuccessfully tried to use the card as soon as they received it. They may have become discouraged and never attempted to provide their census data.

The difference in the effect of the incentive in the CATI and Internet conditions may be attributable to the fact that CATI is easier than the Internet to access for most households. A respondent simply has to pick up a telephone to use CATI. By contrast, many respondents do not have convenient, speedy, and reliable Internet access. The National Telecommunications and Information Administration (2000) reported that only about 41.5 percent of all households had Internet access at the time of Census 2000. Therefore, when households were offered an incentive for using an alternative response mode, the offer was more inviting when the alternative mode was CATI, which requires only a telephone, than when the alternative mode was a web survey, which requires Internet access.

The calling card motivated the respondents to use an alternative response mode. However, the differences between the incentive and no-incentive conditions could also be explained as the combined effects of the calling card and the insert upon which it was attached. The mailings to the households in the incentive condition included a cover letter, a census form, and an insert on heavy stock paper holding the calling card. These inserts described the computer-mediated response alternative in large, colorful text. The mailings to the households in the no-incentive condition included only a cover letter and a census form, without any insert. People in the no-incentive households who ignored the cover letter would not have learned that they had the option of responding using a computer-mediated method. The inserts in the mailings to the households in the incentive condition, however, were not easily ignored. Because of these inserts, people in the incentive-condition households may have been more likely than people in no-incentive households to be aware that they could respond by an alternative, computer-mediated mode. Moreover, the inserts may have conveyed a sense that the computer-mediated response alternative was important and desirable.

The design of the RMIE did not include a way to distinguish between the effect of the calling card and the effect of the insert. Nonetheless, the desirability of the calling cards to the respondents can be measured as the extent to which the respondents actually used the cards that they were given. Figure 11 displays the balances on the calling cards as of October 26, 2000, when the respondents had their cards for six to seven months. A total of 889 cards were activated in this study. The calling card company was unable to provide the balances for 27 of those cards; those cards are excluded from Figure 11. When cards were activated, they were credited with $7.50, the value of 30 minutes of domestic calls. Therefore, cards with $7.50 balances were never used. Cards with zero balances were used for calls worth a total of $7.50. Figure 11 shows that just under one-third of the households whose calling cards were activated never used their calling cards. Another ten percent had placed no more than $1.00 worth of calls with the card. Only one-fourth of the respondents had used up all of the card's value.

Calling Card balances as of October 26, 2000 (Figure 11)



These results suggest that many respondents did not consider the card to be very important, and never used it. Of course, some people may carry an unused calling card just for emergencies, and some people may have misplaced their cards. However, the results on Figure 11 do seem to indicate that the calling card may not have been a universally powerful motivator. The insert, rather than the calling card itself, may have played a key role in motivating the respondents to use the alternative response mode.

4.1.2 Effects of response mode on response rates

This section presents the results regarding these questions:

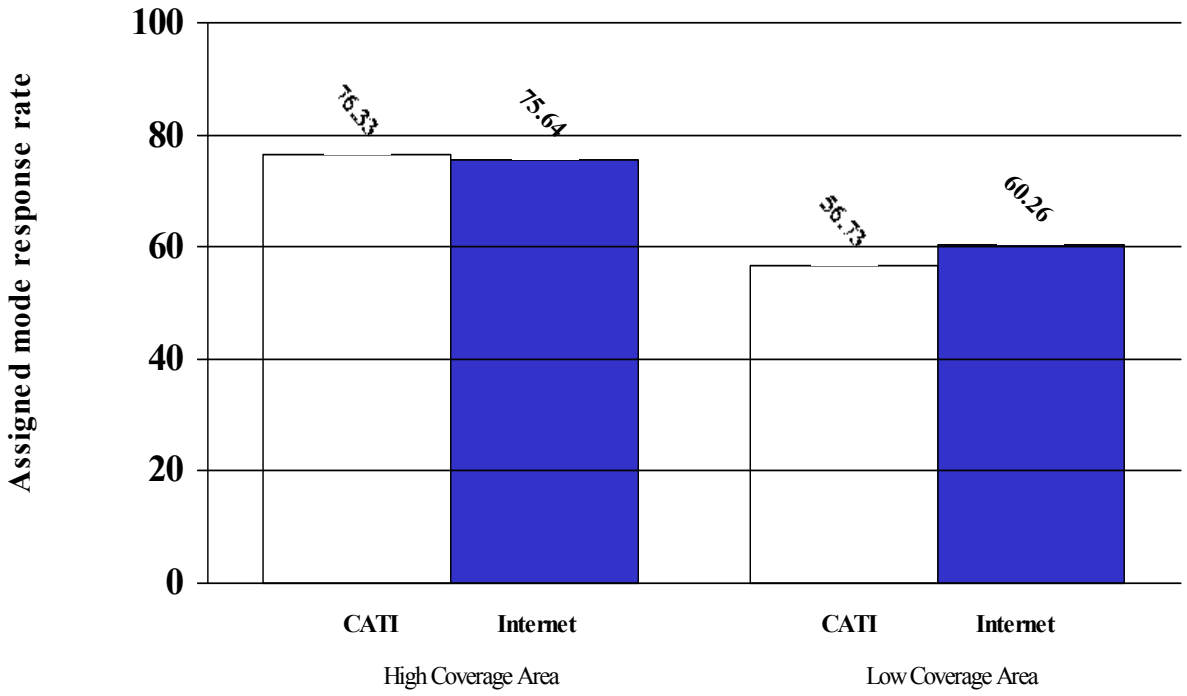
- How did the CATI and the Internet response mode options affect the response rates?
- Did the CATI and Internet response modes have the same impact in the incentive and no-incentive conditions and in the high and low coverage areas?

Effect of Response Mode on the Overall Response Rate. The overall response rate of the CATI panels (panels 1 and 4; 72.66 percent) was not significantly different from that of the Internet panels (panels 3 and 6; 72.70 percent).

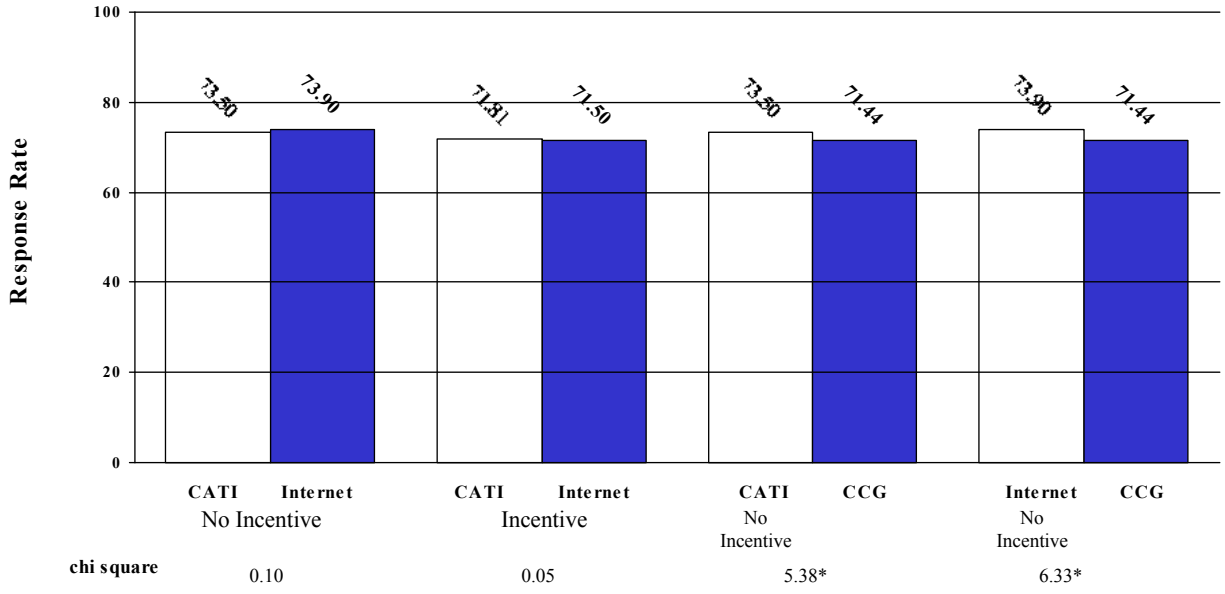
The logistic regression results in Table 1 revealed a significant ($p < .10$) interaction for the overall response rate between the response mode factor and the coverage area. Figure 12 shows that the overall response rate was slightly higher for the CATI panels than for the Internet panels in the high coverage area, but slightly higher for the Internet panels than for the CATI panels in the low coverage area.

Figure 13 shows comparisons between the CATI and Internet panels within the incentive condition and within the no-incentive condition. The figure reveals no difference in the overall response rate between the CATI panel and the Internet panel, regardless of whether or not an incentive was offered.

Overall response rates (Figure 12)



Overall response rates (Figure 13)



***p < .02 (adjusted for multiple comparisons)**

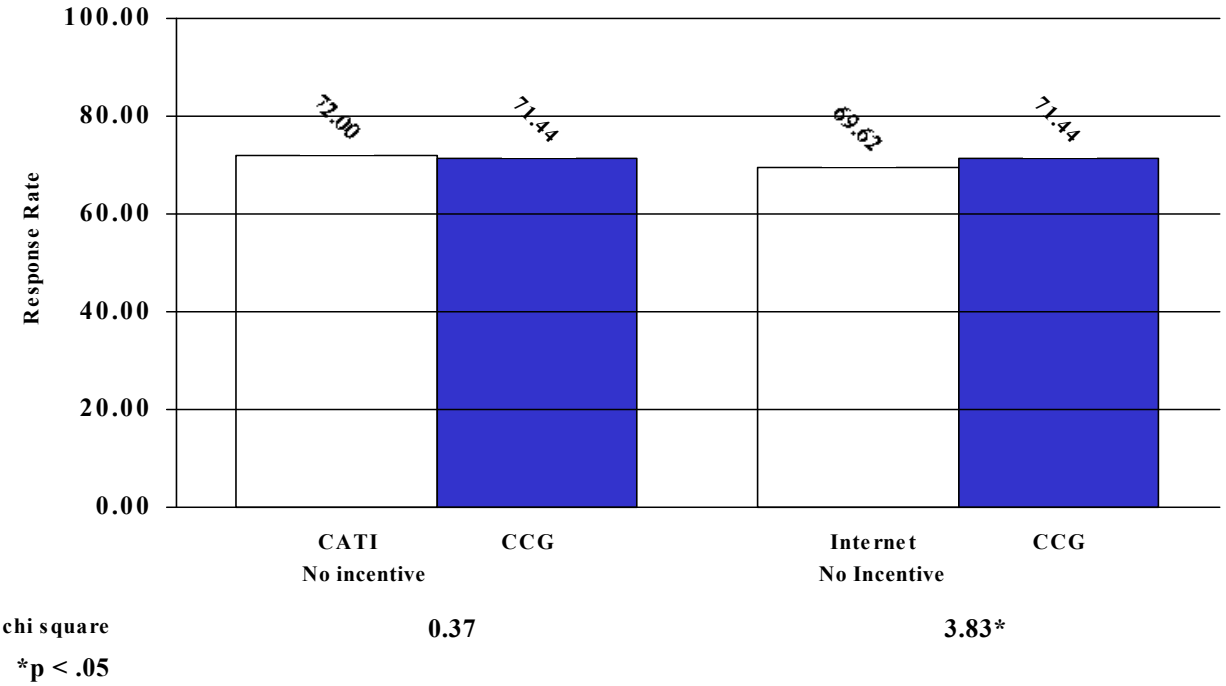
Figure 13 also shows comparisons with CCG. The figure reveals that both the CATI no-incentive panel's and the Internet no-incentive panel's overall response rates were higher than the CCG's. Both of these effects were small – 2.06 percent for the CATI no-incentive panel and 2.46 percent for the Internet no-incentive panel – but they were significant at the $p < .02$ level.

The estimates of significance in the statistical analyses shown on Figure 13 were adjusted for multiple comparisons using a modified Bonferoni procedure (Benjamini & Hochberg, 1995). Since the analyses shown on the figure involved a large number of comparisons, this procedure was necessary to reduce the likelihood that any comparison would reach statistical significance purely by chance.

The finding that the CATI and Internet panels' overall response rates were greater than the CCG's overall response rate suggests that some households that ordinarily would not respond to the census did respond when CATI and the Internet were offered as alternative response modes. The overall response rate went up about the same amount when the CATI alternative was offered and when the Internet alternative was offered. This finding raises a question: Was the increase in the overall response rate due to an increase in responding by mail, or to the additional responding by the alternative response mode?

Figure 14 addresses this question, presenting comparisons of the response rates by the mailed paper forms only. The figure reveals that the mail response rate for the Internet no-incentive panel was lower than the mail response rate for the CCG. The mail response rates did not differ between the CATI no-incentive panel and the CCG. Thus, some of the households that ordinarily would respond by paper forms chose to respond by the Internet instead when it was offered as an alternative. However, not many households that ordinarily would use the paper forms switched to the CATI response mode when it was offered as an alternative.

Response Rate: Paper forms only (Figure 14)



A logistic regression analysis was run that included the CATI no-incentive panel, the Internet no-incentive panel, and the CCG. The outcome variable was the overall response rate. The predictor variables were the response mode (CATI, Internet or CCG), the coverage area (high or low), and all of the response mode by coverage area interaction terms. The full and final models are shown in Table 3. The reference category for the response mode factor was the Internet; the reference category for the coverage area factor was the low coverage area. The results show that no interaction term was statistically significant. The overall response rate differed in the high and low coverage areas, as one would expect. The overall response rate differed between the CCG and the Internet no-incentive panel, as Figure 13 suggests. The difference in the overall response rate between the Internet no-incentive panel and the CATI no-incentive panel was not statistically significant.

Logistic regression analyses of overall response rates: Model with CCG (Table 3)

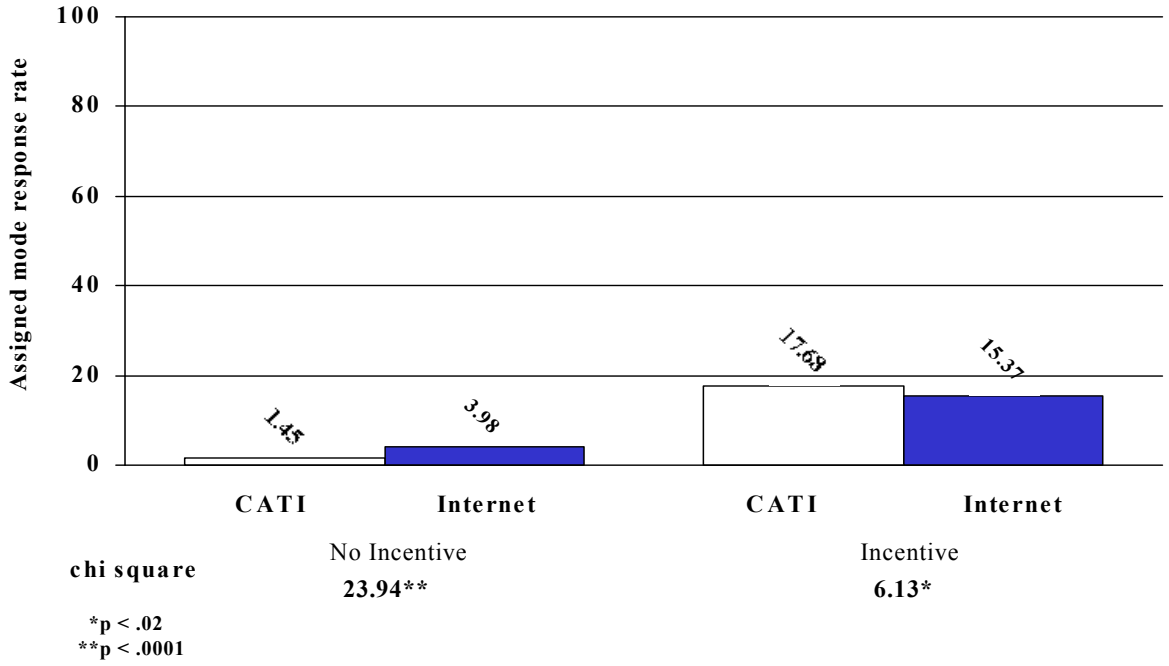
Predictor variable	Full model Coefficient	Standard Error	Final model Coefficient	Standard Error
Intercept	0.46***	(0.094)	0.43***	(0.037)
CATI	-0.18	(0.119)		
CCG	-0.14*	(0.104)	-0.12**	(0.036)
High Coverage Area (HCA)	0.74***	(0.105)	0.77***	(0.034)
CATI x HCA	0.20	(0.135)		
CCG x HCA	0.01	(0.117)		

*p < .05
**p < .002
***p < .001

Effect of Response Mode on the Assigned Mode Response Rates. The assigned mode response rate of the CATI panels (panels 1 and 4; 9.63 percent) was not significantly different from that of the Internet panels (panels 3 and 6; 9.76 percent; chi squared = 0.06, df =1, n.s.).

Figure 15 shows the results of comparisons among the four panels' assigned mode response rates. The figure shows that in the incentive condition, the CATI panel's assigned mode response rate was greater than the Internet panel's. However, in the no-incentive condition, the Internet panel's assigned mode response rate was higher than the CATI panel's. These results are consistent with the significant CATI by incentive interaction term in Table 2.

Assigned mode response rates (Figure 15)



Discussion: The Effects of Mode on Response Rates. The results suggest the following:

- Both the CATI and Internet no-incentive panels had an overall response rate that was higher than the CCG's (see Figure 13). Thus, when an alternative response mode was offered, some households that ordinarily would not respond decided to respond.

- The Internet no-incentive panel was significantly less likely to respond by mail than was the CCG. Thus, when the Internet was offered as an alternate response mode without the incentive, some of households that ordinarily would have responded on the paper forms chose to use the Internet instead.

- The CATI no-incentive panel had about the same mail response rate as the CCG. Thus, when CATI was offered as an alternate response mode without the incentive, not many households that wished to respond to the census on paper forms switched to the CATI mode.

- When no incentive was offered, the Internet panel's assigned mode response rate was higher than the CATI panel's.

- When an incentive was offered, the CATI panel's assigned mode response rate was higher than the Internet panel's.

The results suggest that the two alternate response modes had somewhat different effects. When the Internet was offered without the incentive, some households that ordinarily would have not responded to the census at all did respond, using the Internet. Also, some households that ordinarily would have responded on the paper forms used the Internet instead. When the CATI was offered without the incentive, some households that ordinarily would have not responded at all used CATI to respond. However, not many households that were willing to respond on the paper forms used CATI instead. In summary, both the Internet and CATI were appealing to a small but statistically significant proportion of households that avoid the paper forms. However, when no incentive was offered, only the Internet was appealing to a significant proportion of households that would ordinarily use the paper forms. As a result, when no incentive was offered, the assigned mode response rate was higher in the Internet condition than in the CATI condition.

When an incentive was offered, however, the reverse held: the assigned mode response rate was higher in the CATI condition than in the Internet condition. This result suggests that the incentive motivates respondents to use CATI more than it motivates them to use the Internet.

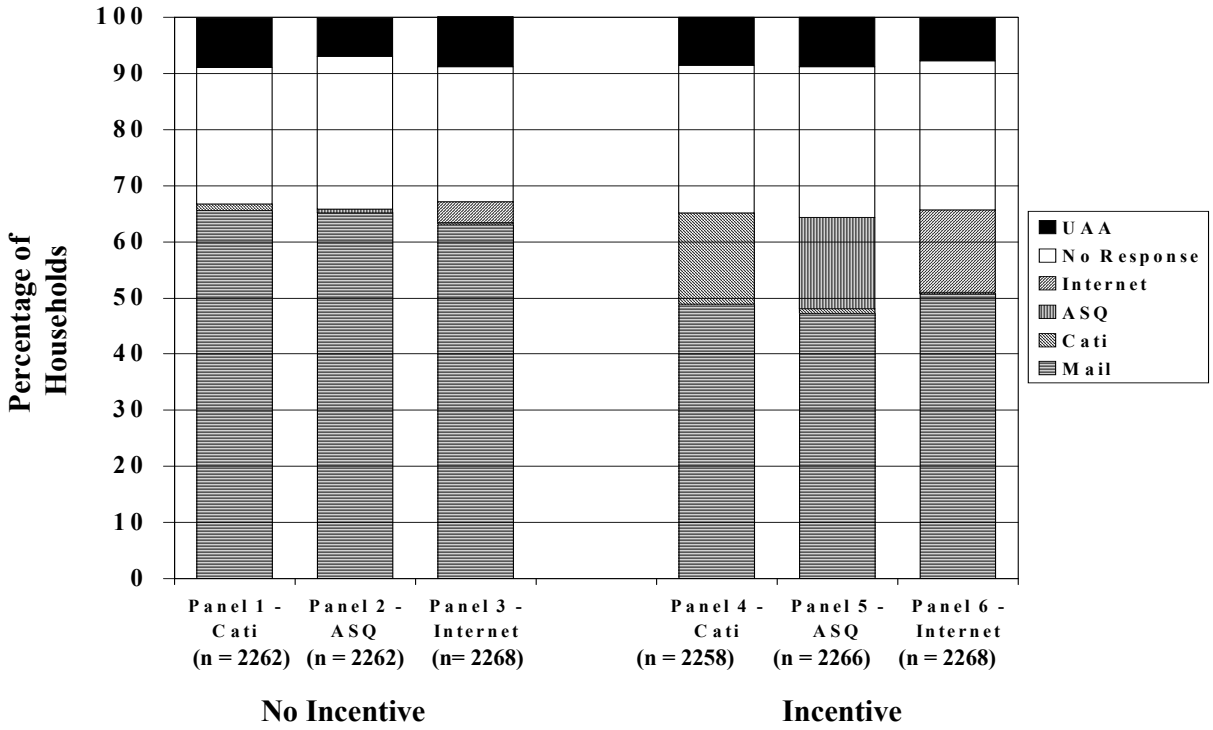
The reason for these results may pertain to the fact that when no incentive was offered, respondents had to read the cover letter to learn about the alternative response mode. Respondents who did not take the trouble to read the cover letter were not aware of the alternative response mode. However, when an incentive was offered, the colorful inserts unmistakably advertised the alternative response modes. Perhaps, the respondents in the no-incentive condition who bothered to read the cover letters were especially likely to have access to the Internet. When these respondents read about the Internet alternative, they used it instead of the paper forms.

However, when an incentive was offered, many more households likely read the colorful inserts and became aware of the alternative response mode. Since more households have telephones than have Internet access, more were likely to use CATI than the Internet.

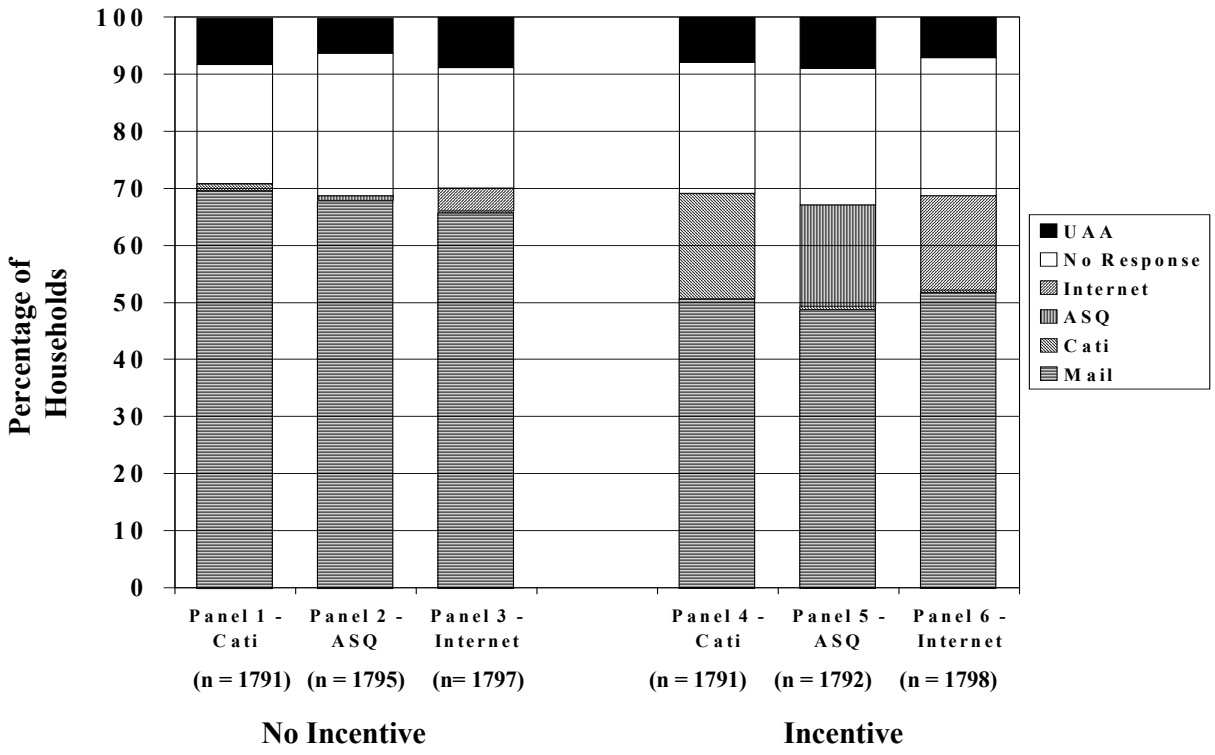
4.2 Effects of the Experimental Conditions on Response Rates: Six Panels, Subnational Sample

These analyses involve all six panels, but exclude households from the ten problematic states and three problematic ZIP Code areas. The responses of the six panels are shown in Figure 16, and separately for the high and low coverage areas in Figures 17 and 18.

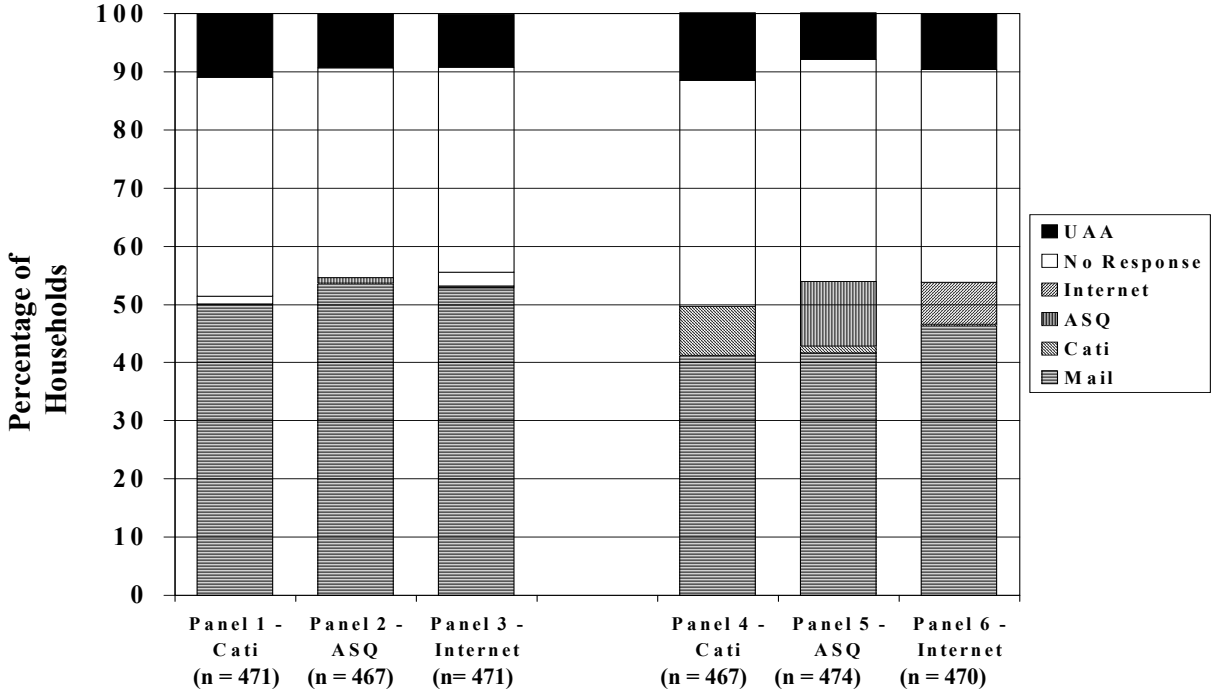
Response to the 2000 Census (Figure 16)



Response to the 2000 Census: High coverage area only (Figure 17)



Response to the 2000 Census: Low coverage area only (Figure 18)



4.2.1 The effects of the incentive on response rates

This section addresses these questions:

- What was the effect of the incentive on the overall and assigned mode response rates?
- Did this effect vary by response mode and by coverage area?

Effects of Incentive on the Overall Response Rate. The overall response rates of the no-incentive panels (panels 1, 2, and 3; 72.55 percent) and the incentive panels (panels 4, 5, and 6; 71.01 percent) were not significantly different (chi square = 2.49, df = 1, n.s.).

A logistic regression analysis was carried out to reveal any significant interactions between the incentive and the two other factors—response mode and coverage area. Table 4 displays the results of that logistic regression analysis. For this regression model, the reference categories were the Internet (for the response mode factor), low coverage area (for the coverage area factor) and no incentive (for the incentive factor).

Table 4 reveals that the incentive factor did not attain statistical significance either by itself or in any interaction.

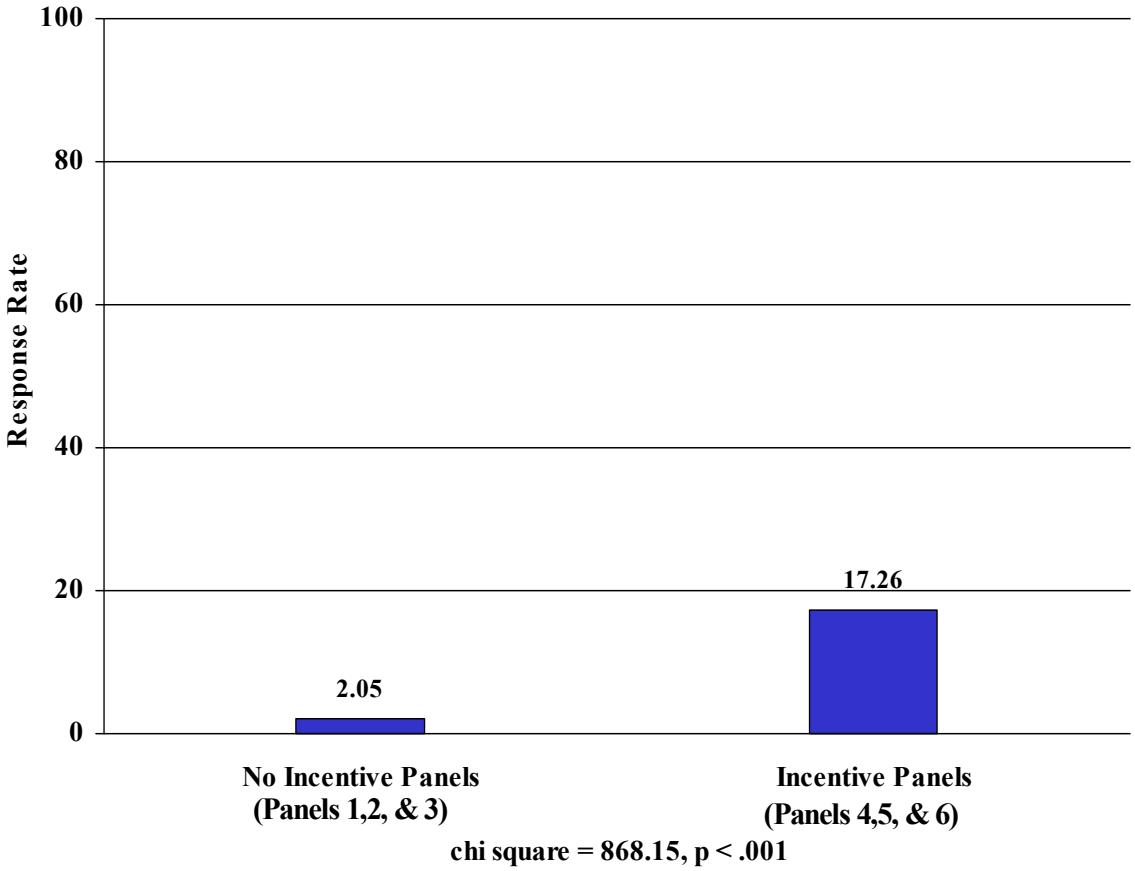
Logistic regression analyses of overall response rates (Table 4)

Predictor variable	Full model Coefficient (Standard Error)		Final model Coefficient (Standard Error)	
Intercept	0.46***	(0.097)	0.44***	(0.051)
CATI	-0.14	(0.121)	-0.16	(0.089)
ASQ	-0.04	(0.124)	-0.08*	(0.043)
Incentive	-0.07	(0.157)		
High Coverage Area (HCA)	0.74***	(0.112)	0.67***	(0.050)
CATI x Incentive	0.01	(0.172)		
ASQ x Incentive	-0.01	(0.155)		
CATI x HCA	0.17	(0.152)	0.21**	(0.092)
ASQ x HCA	-0.14	(0.154)		
Incentive x HCA	-0.08	(0.158)		
CATI x Incentive x HCA	0.02	(0.218)		
ASQ x Incentive x HCA	0.18	(0.176)		

*p < .10
**p < .05
***p < .001

Effect of the Incentive on the Assigned Mode Response Rates. Figure 19 reveals that the incentive was associated with a large increase on the assigned mode response rate.

Assigned mode response rate: combined panels (Figure 19)



A logistic regression analysis was carried out. The results are shown in Table 5.

Logistic regression analyses of the assigned mode response rates (Table 5)

Predictor variable	Full model Coefficient (Standard Error)		Final model Coefficient (Standard Error)	
Intercept	-3.55***	(0.330)	-3.36***	(0.233)
CATI	-0.69*	(0.523)	-1.11***	(0.276)
ASQ	-1.11**	(0.605)	-1.64***	(0.267)
Incentive	1.10***	(0.406)	1.03***	(0.260)
High Coverage Area (HCA)	0.45	(0.372)	0.22***	(0.255)
CATI x Incentive	0.87**	(0.551)	1.26***	(0.287)
ASQ x Incentive	1.53***	(0.688)	1.79***	(0.276)
CATI x HCA	-0.52	(0.564)		
ASQ x HCA	-0.65	(0.719)		
Incentive x HCA	0.47*	(0.449)	0.56*	(0.263)
CATI x Incentive x HCA	0.48	(0.605)		
ASQ x Incentive x HCA	0.34	(0.838)		

*p < .05
**p < .002
***p < .001

Table 5 (final model) reveals a significant interaction between the incentive factor and the response mode factor. The difference between the incentive and no-incentive conditions was greater for the ASQ and CATI response mode than it was for the Internet response mode.

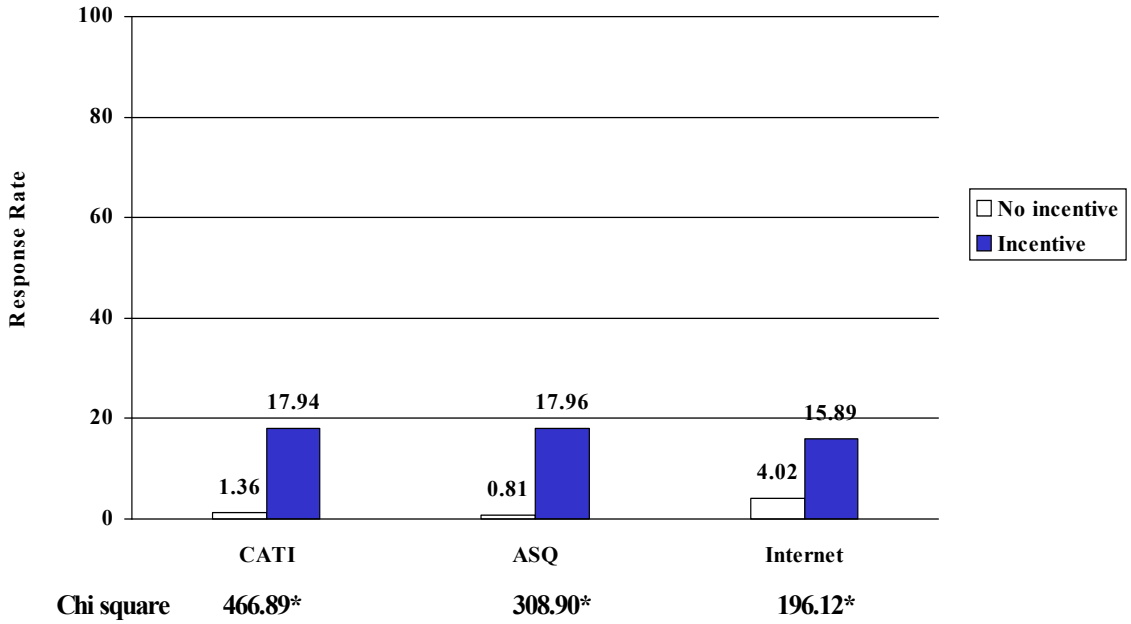
Chi square analyses were carried out to illustrate the manner in which the incentive affected the assigned mode response rates. Figure 20 shows that the assigned mode response rates in the incentive households were higher than those in the no-incentive households, regardless of whether the households were in the CATI, ASQ, or Internet response mode conditions. According to the logistic regression, this difference between the incentive and no incentive condition was larger for the CATI and ASQ conditions than it was for the Internet condition.

Table 5 also reveals a significant interaction between the incentive factor and the coverage area factor. The difference between the incentive and no-incentive conditions was greater in the high coverage area than in the low coverage area. Figure 21 shows that the incentive increased the assigned mode response rate, regardless of whether the households were in the high or low coverage area.

Discussion: The Effect of the Incentive on Response Rates. The effect of the incentive in the analyses involving all six panels and a subnational sample can be summarized as follows:

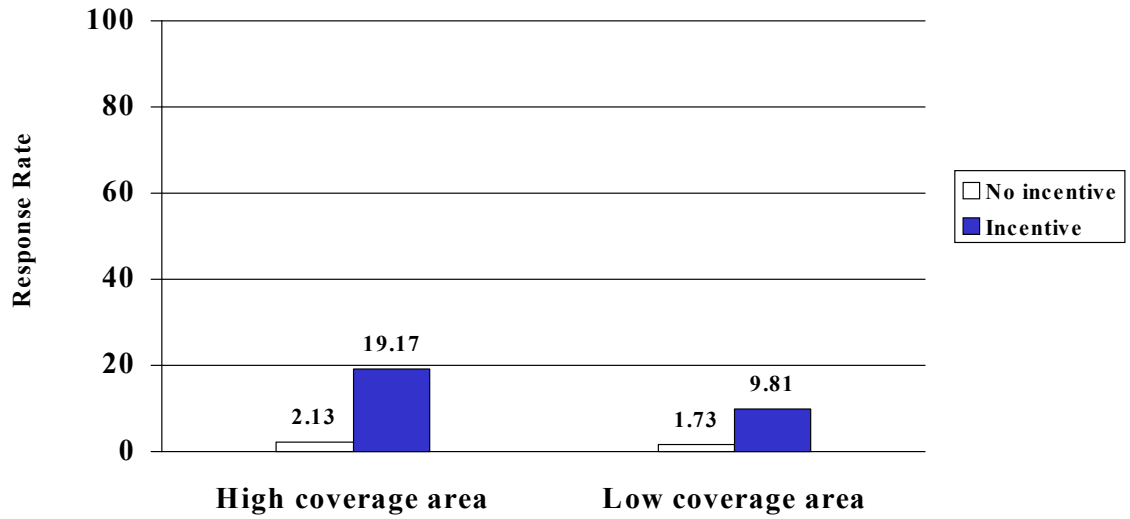
- The incentive offered to the households for responding via an alternative, computer-mediated response mode had no significant effects on the overall response rate.

Assigned mode response rate by panel (Figure 20)



*p < .001

Assigned mode response rate: Incentive by coverage area interaction (Figure 21)



Chi square	824.47*	86.51*

*p < .001

- The incentive increased the likelihood that the households would choose the alternative response mode.
- The incentive increased the assigned mode response rate most for the ASQ and CATI response modes, and least for the Internet response mode.
- The incentive increased the assigned mode response rate more for the high coverage area than for the low coverage area.

As mentioned in the Limitations section above, the possibility exists that the problems with the Panel 2 mailout were not entirely eliminated by excluding ten states and three ZIP Code areas from the analyses. The panel 2 overall and assigned mode response rates may have been depressed by some continuing problem with the data for this panel.

The effect of the incentive in depressing the overall response that was apparent in the earlier analyses was not apparent in these analyses. This change is attributable to the depressed overall response rate in panel 2. This depressed response rate may be caused by a lingering data problem or by some other cause that is difficult to identify. As a result of the depressed overall response rate in panel 2, the mean overall response rate of the no-incentive panels (panels 1, 2, and 3) was not different from the mean overall response rate of the incentive panels (panels 4, 5, and 6).

The incentive increased the assigned mode response rate for all of the alternative response modes. This effect was smallest for the Internet condition, largely because so many households used the Internet even when no incentive was offered. This finding replicates the earlier analyses using the national sample. These results suggest that the Internet has some inherent appeal to respondents, even without the offer of an incentive. The telephone-based response modes, CATI and ASQ, did not have as much inherent appeal.

The incentive increased the assigned mode response rate more for the high coverage area households than for the low coverage area households. This result was not apparent in the earlier analyses involving the national sample. The reason for the finding is not obvious. Perhaps the high coverage area households valued the calling card more. Perhaps they were more amenable to the idea that they would have to use the alternative response mode for the calling card to have value.

4.2.2 Effects of the response mode on response rates

This section addresses the following questions for six panels and a subnational sample:

- How did the CATI, ASQ, and the Internet response mode options affect the response rates?

- Did the CATI, ASQ, and Internet response modes have the same impact in the incentive and no-incentive conditions and in the high and low coverage areas?

Effect of Response Mode on the Overall Response Rate. The overall response rates of the CATI panels (72.33 percent), ASQ panels (70.67 percent) and Internet panels (72.35 percent) were not significantly different (chi square = 4.32, df = 2, n.s.).

The logistic regression analysis shown in Table 4 reveals a significant interaction between the response mode factor and the coverage area factor. Respondents in the high coverage area were more likely to use CATI than the Internet. Chi square analyses were run to further illustrate the relationship between the response mode factor and the coverage area factor. The results suggested that the overall response rates differed among the three response mode conditions in the high coverage area (chi square = 7.05, df = 2, $p < .03$) but not in the low coverage area (chi square = 2.30, df = 2, n.s.). For high coverage area households, the overall response rate was lower in the ASQ condition (73.6 percent) than in either the CATI condition (76.2 percent, chi square = 6.23, df = 1, $p < .02$), or the Internet condition (75.4 percent, chi square = 3.21, df = 1, $p < .08$). No significant difference was found in the high coverage area households between the overall response rates in the CATI and Internet conditions (76.2 percent and 75.4 percent respectively, chi square = 0.64, df = 1, n.s.).

A logistic regression analysis was run that included the CATI no-incentive panel, the ASQ-no incentive panel, the Internet no-incentive panel, and the CCG. This analysis was similar to the one shown in Table 3, which involved the national sample and four panels. The outcome variable was a response indicator. The predictor variables were the response mode, the coverage area, and all of the interaction terms. The full and final models are shown in Table 6. None of the interaction terms was statistically significant.

To further illustrate the pattern across response modes, chi square analyses compared the overall response rates of the CCG (71.1 percent) with those of the CATI no-incentive (72.33 percent), ASQ no-incentive (70.67 percent), and Internet no-incentive (72.35 percent) panels. The overall response rate of the CCG was lower than that of the CATI no-incentive panel (chi square = 2.89, df = 1, $p < .09$), and the Internet no-incentive panel (chi square = 4.29, $p < .04$). The overall

response rates of the CCG and the ASQ no-incentive panel did not differ (chi square = 0.26, df = 1, n.s.).

Logistic regression analyses of overall response rates: Model with CCG (Table 6)

Predictor variable	Full model Coefficient (Standard Error)		Final model Coefficient (Standard Error)	
Intercept	0.46***	(0.097)	0.44***	(0.041)
CATI	-0.14	(0.121)		
ASQ	-0.04	(0.124)	-0.14*	(0.055)
CCG	-0.13*	(0.108)	-0.12**	(0.041)
High Coverage Area (HCA)	0.74***	(0.112)	0.75***	(0.034)
CATI x HCA	0.17	(0.152)		
ASQ x HCA	-0.14	(0.154)		
CCG x HCA	0.01	(0.124)		

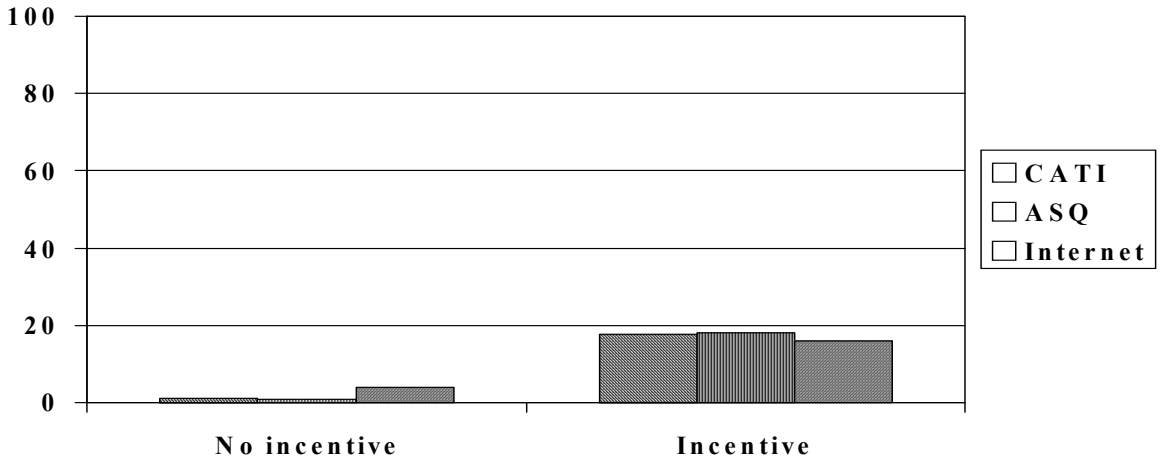
*p < .05
**p < .005
***p < .001

Effect of the Response Mode on the Assigned Mode Response Rate. A three by two chi square test compared the assigned mode response rates of the CATI panels (9.65 percent), ASQ panels (9.30 percent) and Internet panels (10.0 percent). The differences were not significantly different (chi square = 1.53, df = 2, n.s.).

Table 5 reveals a significant interaction between the incentive factor and the response mode factor. This interaction suggests that the incentive increased the assigned mode response rate in the CATI and ASQ conditions more than in the Internet condition. To further illustrate the nature of this interaction, chi square tests were run between the incentive factor and the response mode factor, as shown in Figure 22. The results suggest that in the no-incentive condition, the Internet panel had the *greatest* assigned mode response rate (versus the CATI panel, chi square = 27.09, df = 1, p < .001; versus the ASQ panel, chi square = 61.01, df = 1, p < .001). The assigned mode response rates of the CATI and ASQ panels did not differ (chi square = 2.64, df = 1, n.s.).

For the incentive condition, Internet panel had the *lowest* assigned mode response rate (versus the CATI panel, chi square = 4.77, p < .03; versus the ASQ panel, chi square = 2.98, p < .09). Again the assigned mode response rates of the CATI and ASQ panels did not differ (chi square = 0.00, df = 1, n.s.).

Assigned mode response rate by panel (Figure 22)



Note–Chi square values are presented in the text.

Discussion: Effect of the Response Mode on Response Rates. The effect of the incentive in the analyses involving all six panels and a subnational sample can be summarized as follows:

- The overall response rate did not differ across the CATI, ASQ, and Internet conditions.
- In the high coverage area, the overall response rate in the ASQ condition was lower than the overall response rate in the CATI or Internet conditions.
- The CATI no-incentive and the Internet no-incentive panels had a higher overall response rate than the CCG.
- The CCG's overall response rate was not significantly different from that of the ASQ no-incentive panel.
- In the no-incentive condition, the Internet panel had the *greatest* assigned mode response rate.
- In the incentive condition, the Internet panel had the *lowest* assigned mode response rate.

These results replicate those for the analyses involving the national sample and only four panels. The one new finding is the lack of a difference between the CCG's and ASQ no-incentive panel's overall response rate. Thus, offering the Internet and CATI alternatives appeared to boost the overall response rate, relative to the CCG, but offering the ASQ did not. The assigned mode response rate for the ASQ no-incentive panel was only 0.81 percent. Perhaps lingering mailout problems with the ASQ no-incentive panel depressed its overall response rate. Or, perhaps a number of respondents hung up when they heard the ASQ system, even before they entered their Census ID numbers. The ASQ system did not capture any data about such hang ups. Conceivably,

some respondents hung up this way and then never provided their census data using any response mode at all.

4.3 Item Nonresponse Rates by Mode of Response

Computer-assisted modes of data collection may yield better quality data than paper-and-pencil data collection modes can provide. Computer-assisted data collection methods can apply automated skips and edits to reduce the amount of missing data. In the present study, the CATI and ASQ modes employed automated skips and edits, but the Internet mode did not. This section presents comparisons of the amount of missing data across the different response modes in RMIE. For these analysis, "response mode" refers to the mode that the respondent actually used to provide the data, not necessarily the mode that was assigned in the experiment. For example, if a respondent called into the ASQ and rolled over to a CATI interviewer, the questionnaire would be considered to be in the CATI mode. If a respondent assigned to the CATI condition provided data using the Internet, the questionnaire would be considered to be in the Internet mode.

Analysis. The questions on the census short form were divided into household-level items (i.e., questions about the entire household) and person-level items (i.e., questions about the individuals who comprise the household). There were two household-level items: population count and home ownership. There were eight person-level items: name, relation, gender, year of birth, month and day of birth, age, Hispanic background, and race. Of course, the relation item did not apply to person 1 in the household; the relation was always "self."

Table 7 shows the item nonresponse rates for the two household-level and eight person-level items on the census short form. The table displays the item nonresponse rates for just the first six persons in the household, because the paper form and the Internet survey collected these data for only six persons, even if more than six persons resided in the household. (The ASQ and the CATI survey collected data from all persons in the household, without limit). No statistical tests were performed on these data.

The highest item nonresponse rates occurred when the data were collected using the ASQ, up to 11.8 percent (e.g., person 1 – race) (Table 7). Much lower rates occurred when the data were collected by the other modes. Among these other modes, the mail has the highest item nonresponse rates, with the Internet and CATI having the lowest rates.

Percent item nonresponse by questionnaire item and mode of response (Table 7)

Items	Mail(n=8575)		ASQ (n=220)		CATI (n=700)		Internet (n=457)	
	%	see	%	see	%	see	%	see
Household Items								
Pop Count	2.2	.15	NA	-	0	-	0	0
Ownership	2.0	.16	4.1	1.17	0	-	0	0
Persons Items								
Person 1								
Name	0.8	.10	0.9	.63	0	-	0.7	.37
Gender	0.7	.09	6.8	1.83	0	-	0.7	.37
Year of birth	1.0	.10	10.0	2.11	.1	.14	0.9	.43
Day & month	1.3	.11	10.5	2.03	.1	.14	1.3	.51
Age	1.2	.14	8.2	1.98	0	-	0.9	.43
Hispanic	3.1	.17	6.4	1.81	.1	.14	1.1	.47
Race	2.0	.16	11.8	2.06	.1	.14	0.4	.31
Person 2								
Name	0.6	.08	0	-	0	-	0.7	.37
Relation	0.6	.08	6.8	1.57	0	-	0.4	.31
Gender	0.5	.08	5.5	1.46	0	-	0.7	.37
Year of birth	1.3	.12	6.4	1.55	.4	.24	0.9	.53
Day & month	1.3	.12	6.4	1.55	.4	.24	0.9	.53
Age	1.2	.14	5.9	1.50	0	-	0.7	.37
Hispanic	2.3	.16	5.9	1.42	.1	.14	1.1	.47
Race	2.4	.16	7.7	1.56	.1	.14	0.9	.42
Person 3								
Name	0.4	.07	0	-	0	-	0.2	.22
Relation	0.3	.06	5.5	1.43	0	-	0.2	.22
Gender	0.7	.08	4.5	1.37	0	-	0.2	.22
Year of birth	0.8	.10	5.0	1.42	.6	.28	0.7	.37
Day & month	0.8	.10	5.0	1.42	.6	.28	0.7	.37
Age	0.8	.11	4.5	1.37	.14	.14	0.4	.31
Hispanic	1.4	.12	4.5	1.37	0	-	1.3	.53
Race	1.8	.15	5.9	1.42	0	-	0.7	.38
Person 4								
Name	0.2	.04	0	-	0	-	0	-
Relation	0.3	.04	3.6	1.09	0	-	0	-
Gender	0.7	.11	3.6	1.12	0	-	0.2	.22
Year of birth	0.5	.08	3.6	1.14	.1	.14	0.2	.22
Day & month	0.6	.08	3.6	1.14	.1	.14	0.2	.22
Age	0.8	.08	3.6	1.15	.1	.14	0.4	.31
Hispanic	0.8	.11	3.2	1.08	0	-	0.2	.22
Race	1.4	.13	3.2	1.08	0	-	0	-
Person 5								
Name	0.0	.02	0	-	0	-	0	-
Relation	0.1	.03	3.2	1.09	0	-	0	-
Gender	0.3	.06	2.7	1.02	0	-	0.2	.22
Year of birth	0.3	.07	2.7	1.02	0	-	0	-
Day & month	0.3	.07	2.7	1.02	0	-	0	-

Percent item nonresponse by questionnaire item and mode of response (Table 7) (continued)

Items	Mail(n=8575)		ASQ (n=220)		CATI (n=700)		Internet (n=457)	
	%	see	%	see	%	see	%	see
Age	0.4	.07	3.2	1.24	0	-	0	-
Hispanic	0.4	.07	2.7	1.02	0	-	0	-
Race	0.7	.10	2.7	1.02	0	-	0	-
Person 6								
Name	2.3	.18	0.9	.65	1.3	.45	2.4	.65
Relation	0.1	.03	0.4	.46	0	-	0	-
Gender	0.1	.04	0.4	.46	0	-	0	-
Year of birth	0.1	.03	0.4	.46	0.1	.14	0	-
Day & month	0.1	.03	0.4	.46	.14	.14	0	-
Age	0.2	.04	0.4	.46	0	-	0	-
Hispanic	0.1	.03	0.4	.46	0	-	0.2	.22
Race	0.2	.04	0.4	.46	0	-	0	-

Note: ASQ = Automated Spoken Questionnaire; CATI = computer-assisted telephone interviewing; n = unweighted number of questionnaires with missing item. NA = Item was not collected directly by the ASQ, but was computed using the number of names that were reported by the respondent at the beginning of the questionnaire. SEE = standard error of the estimate, which is undefined and marked by "-" when the percentage is zero.

Much of the missing data on the ASQ was due to respondents' stopping the interview entirely before it was complete. A total of 16 of the 220 ASQ respondents broke off the interview in this way. For example, of the 22 respondents where year of birth for the first person in the household was missing, 13 hung up either at or before this point in the interview. A similar pattern occurred for the other items in the ASQ. While it is difficult to determine exactly why any single individual hung up, the respondents' recorded comments suggested that an important reason was the length of the interview. Several respondents commented that "...this was taking too long..." Similarly, in some voice files there is a tone or manner of speaking by the respondent that could be interpreted as frustration with the pace of the interview. Interestingly, respondents who chose to hang up frequently did so at the beginning of the interview, either during the household questions or within the first few items of the first person in the household (e.g., prior to asking for the age or date of birth). Three of the 16 people who hung up before the end of the interview did so during the race question for the first person, which is the last question about the first person.

Some of the missing data for the ASQ were not due to the respondents' hanging up the telephone. Some missing data were isolated to particular items (e.g., year of birth is missing, but all other items are present). These missing data were of two types. In one type, the respondent seemed to be trying to speak but was cut off by the program. One gets the impression listening to these files (not grounded in any direct data) that respondents had not quite decided what they needed to report and ran out of time before being cut off by the time limit set on the computer (two seconds). For example, on a number of the missing data files, one can hear the respondent inhaling as if to speak but then being cut off before actually being able to do so. Of the 22 cases that are missing on year of birth for the first person in the household, 5 cases appear to be this type of missing data.

A second type of missing data is silence, without any hint that the respondent was trying to answer the question. This primarily occurred for the race question, where 10 of the 26 missing cases for this item were this type of response that was, in part, a result of the way the race question was structured and programmed. For these items, the respondent was first asked:

A. "Do you belong to one or more of the races printed under question 6 of the short form questionnaire?"

A "yes" response was followed up with:

B. Please say the name of the race or races you belong to with a short pause between each name.

A "no" response was followed up with:

C. Please say the name of the other race or races you belong to with a short pause between each name.

If the computer could not recognize the response to question A above, or if there was no response, the respondent was skipped to a series of "yes/no" items that asked if the person was or was not in a particular race group (e.g., White, American Indian, Asian Indian).

In eight of the ten missing cases on race, the computer interpreted nonhuman noise as an "utterance" and did not skip the respondent to the list of "yes/no" questions. Consequently, these individuals never provided responses on their racial classification. In the other two instances, the computer did not detect a response at item B or C, and the followup questions were asked. One of these two respondents answered "no" to the entire list of race items at this point. This individual was Hispanic and may have considered the race questions redundant with the Hispanic question (which preceded the race item). In the other instance, the respondent did not respond to the item asking if he/she was "white." This person subsequently reported "no" to all the other race-specific questions in this series.

Among ASQ respondents, for the first person in the household, a total of 189 answered question A above with "yes," and provided their races from the list provided. A total of 28 answered question A with "no," and provided their races, although the races were not on the list provided. A total of 18 did not answer question A and went on to answer the individual race questions. For all other persons in the household, 258 answered question A with "yes" and provided the races from the list, 29 answered with "no" and provided the races although they were not on the list, and ten answered the individual questions about race.

The other automated modes (CATI and Internet) had less missing data than the mail questionnaires did on several items, although the amount of missing data across all three of these modes was relatively low. The largest differences are for the Hispanic and race information. For the mail survey 3.1 percent of the person 1 data are missing on the Hispanic question, while only .3 percent and 1.1 percent are missing on the CATI and Internet, respectively. This pattern of missing data for Hispanic and race information occurs for the person 2 data as well.

Discussion. The amount of missing data for the ASQ has important implications for the feasibility of this mode for the decennial census. Up to ten percent of the responses were missing for some

of the demographic items (e.g., age, date of birth), posing significant data quality problems. As noted above, there were two different sources of the missing information. A large proportion was due to ASQ respondents' hanging up before the end of the interview. Most of these hang-ups occurred early in the interview. Some comments from respondents indicated impatience with the pace of the interview. This reaction may have been exacerbated by the type of information that was collected at the beginning of the interview, when the respondents were asked to enter their 22-digit identification numbers and telephone numbers with touch-tone buttons, and to say and spell the names of everyone in the household. These tasks, along with the speed with which the questions were administered, may have played a role in the respondents' decision to terminate the interview prematurely. Future ASQ designers will need to streamline the beginning of the interview so that respondents are not frustrated with the pace. In addition, it may also be desirable to inform the respondents before they call that they would be providing data to a computer, rather than a person. The respondents would then know what to expect and be less likely to feel frustrated.

Some of the missing data in the ASQ mode may be attributable to problems respondents encountered providing data within the time constraints allotted by the computer program. The system was programmed to repeat the question when it encountered two seconds of silence. Even given this repetition, respondents sometimes could not report the information for some items. Future ASQs may require a longer wait time after each question. That is, the system should give the respondents more time to begin answering before it repeats the question. A longer wait time has relatively little cost (e.g., it does not increase the length of time to fill out the questionnaire for those that provide answers right away) and could result in capturing data from some of the respondents who, for whatever reason, could not initiate their answers within two seconds.

Another set of issues concern the race questions on the ASQ. A relatively large proportion of the missing data for this item occurred because respondents did not provide a response to the initial open-ended questions (see items A, B and C above). Theoretically, these respondents should have been skipped to the list of "yes/no" questions asking about each particular race. For most of these cases (eight of ten), noise occurred which the system wrongly interpreted as a response. The system therefore skipped the individual "yes/no" items and these respondents never reported their race. Perhaps, these respondents did not respond to the open-ended questions because they could not locate their short form (which was required to answer these items). Even those who could locate their form may not always have been able to check the list of races within the time limit built into the program.

The items about race might be better designed so that respondents are less dependent on having and locating the paper form. One possibility would be to read out the categories to the respondent as part of the question, rather than having respondents read the categories from the paper form. For example, respondents might be asked a question like the following:

"Do you belong to more than one of the following race groups: White, Black, African American..."

(If yes) "How many of these groups do you belong to?"

(If yes) "Could you please tell me one of the race groups you belong to?" (Repeat for as many races as needed).

The CATI and Internet had lower rates of missing data than did the mail questionnaire. It is difficult to draw any conclusions from these patterns, because respondents were not randomly assigned to the mail version. Since the use of the alternative modes was voluntary, certain types of respondents chose to use them, while others types of respondents did not choose to do so. Item nonresponse was relatively rare (e.g., five Internet questionnaires out of approximately 450 were missing a response to the Hispanic item). The results must be replicated within a different context before any firm conclusions can be drawn.

Nonetheless, it is interesting to note that the mail questionnaire had the highest rates of missing data for several items, including home ownership, Hispanic identity, and race. The smaller amount of missing data for the alternative modes may have been due to several features of the computerized systems. As noted in the literature review at the beginning of this report, computer-assisted instruments like the CATI virtually eliminate the possibility that respondents inadvertently skip questions. Moreover, CATI interviewers could probe respondents to ensure they provided answers to every item.

We have no clear explanation why the Internet had less missing data than the mail questionnaire. The Internet questionnaire did not incorporate automated skip patterns, so respondents were free to fill in answers to questions in any particular order (as in the paper form). In that regard, therefore, the Internet and mail questionnaires were very similar. Therefore, no clear conclusions can be drawn from the differences observed between these two modes.

4.4 ASQ Satisfaction Survey – Results

Table 8 and Figures 23 through 34 display the data concerning respondent satisfaction collected by the ASQ for 275 calls that were received in response to the initial and NR phase mailouts. The table and figures show descriptive statistics concerning the call itself, the data provided by the respondent on the census questionnaire, and the respondent's evaluation of the ASQ.

When the speech recognition software attempted to recognize the respondent's utterances in response to the questions on the ASQ, software logged a confidence level for each recognition attempt. For each question, a confidence level threshold had been preset. If the respondent was silent or the speech recognizer's confidence level for a particular recognition was below that threshold, the ASQ repeated the question. After two such attempts, the ASQ either went on to the next question or transferred the call to a CATI operator, depending on the question involved.

The threshold levels for the various questions were determined by trial and error. For most questions, utterances that were not valid responses yielded confidence levels below 70 percent; utterances that were valid responses, even when spoken hoarsely, with accents, or with background noise, yielded confidence levels higher than 70 percent. For that reason, the threshold for most questions was set at 70 percent. For the questions concerning relationship or association, however, a threshold of 70 percent would be too high. These questions asked for the relationship or association between a person in the household and the respondent. These questions had several similar-sounding valid responses, such as, "son" and "son-in-law," or "roomer" and "roommate." Therefore, the threshold for the relationship and association questions was lowered to 60 percent. Also, for the questions asking for the birthdates of the people in the household, the threshold was set at one percent so that all responses would be acceptable.

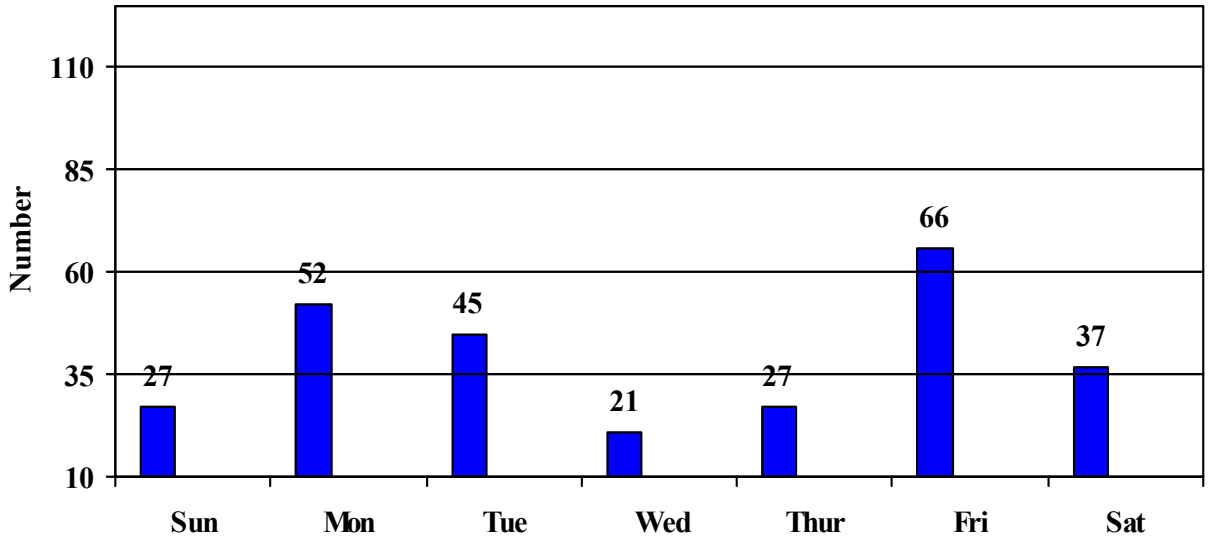
Descriptive statistics for the 275 Automated Spoken Questionnaires (Table 8)

Measure	Number	Percent
ASQ calls, initial mailout	205	
ASQ calls, NR phase mailout	70	
Total ASQ calls	**275**	
Questionnaire items, initial mailout	9,229	
Questionnaire items, NR phase	2,797	
Total questionnaire items	**12,026**	

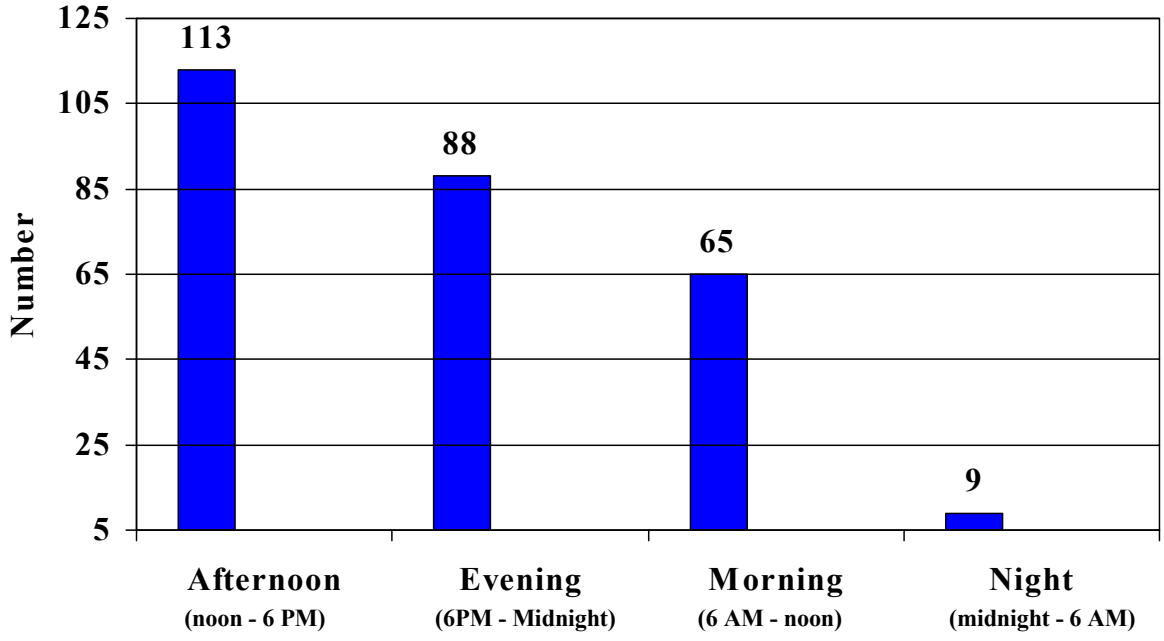
Note: Questionnaire items include both questions to collect Census 2000 data and questions to assess user satisfaction with the ASQ.

Mean (standard deviation) confidence level of Automatic Speech Recognizer	91.75 (6.93)
Mean (standard deviation) confidence level of Automatic Speech Recognizer (calls on or after April 7, 2000 only)	91.73 (6.89)
Mean (standard deviation) number of retries because of a silent reply during call	1.41 (1.70)
Mean (standard deviation) number of retries because of a reply with a subthreshold confidence level during call	0.78 (1.05)
Mean (standard deviation) number of retries in total during call	2.19 (2.21)
Mean (standard deviation) length of entire call (seconds)	902.23 (397.48)
Mean (standard deviation) length of time (seconds) required for: system to pose question; respondent to answer (including two second endpointing); and (if applicable) system to recognize response	16.78 (14.81)
Mean (standard deviation) length of time (seconds) required for: respondent to answer (including two second endpointing); and (if applicable) the system to recognize response	5.18 (3.03)

Day of the week of the call (Figure 23)

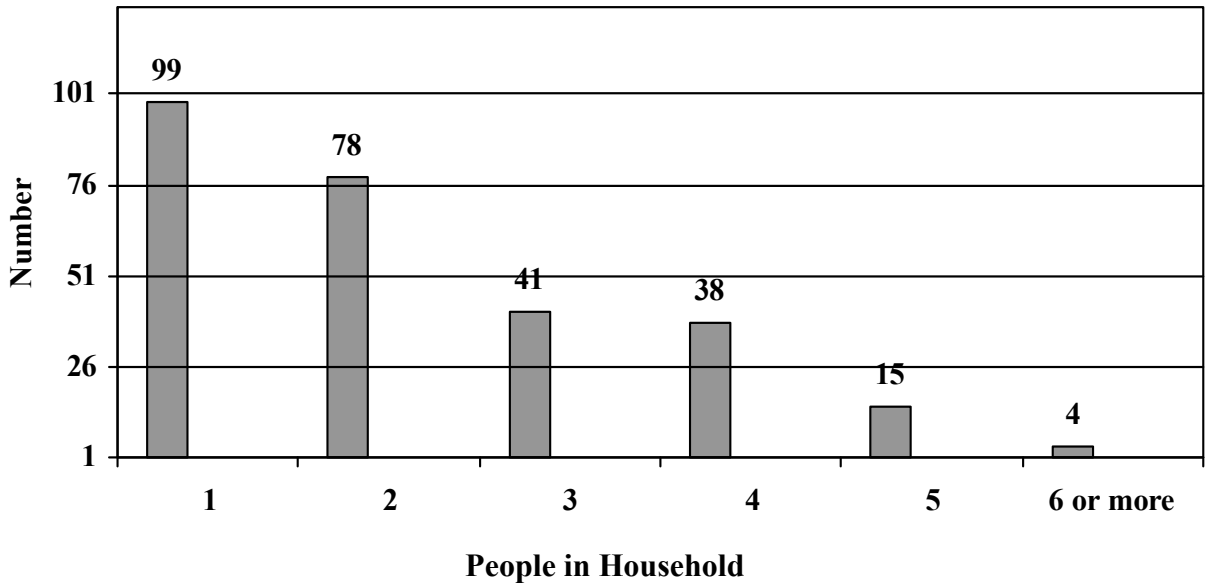


Hour of the call (Figure 24)

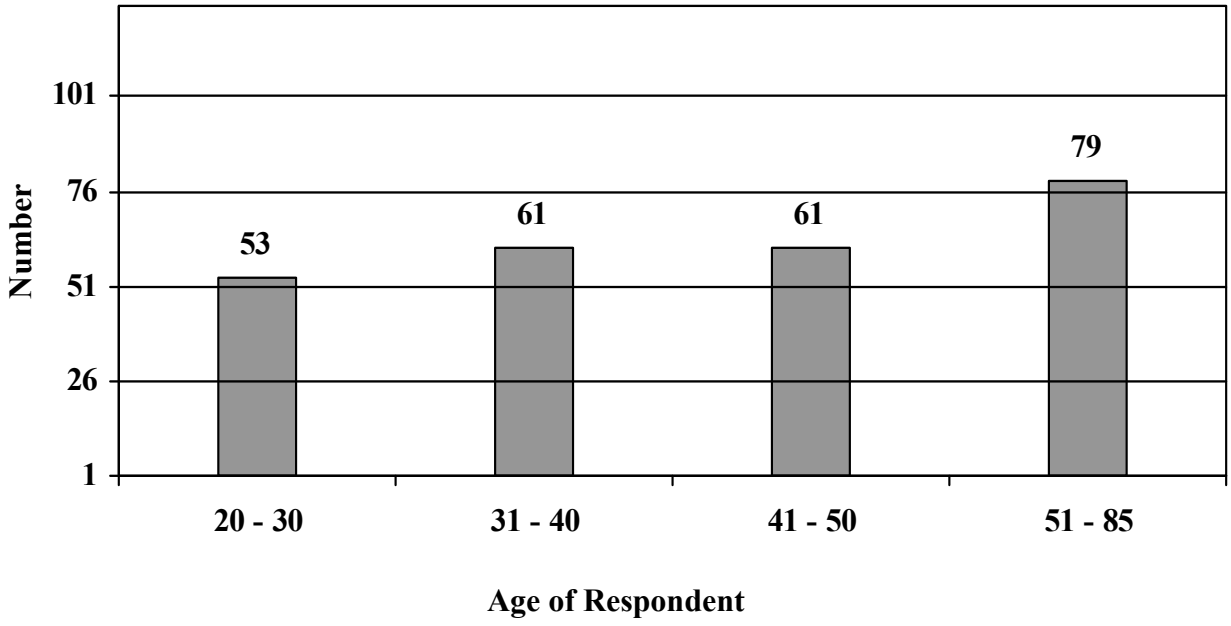


Eastern Time

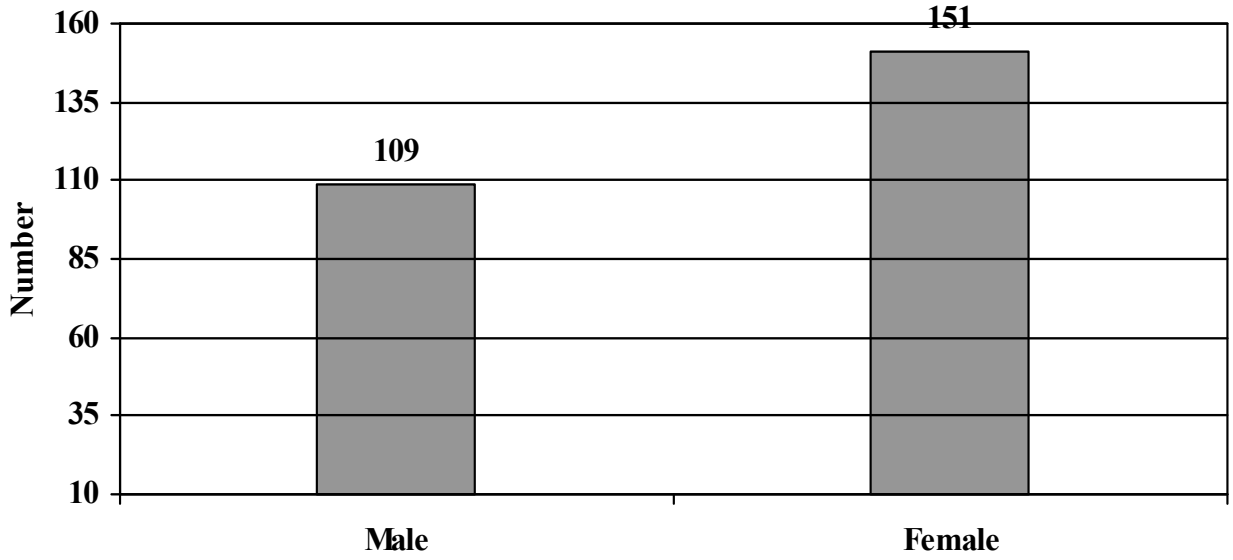
Number of people in household (Figure 25)



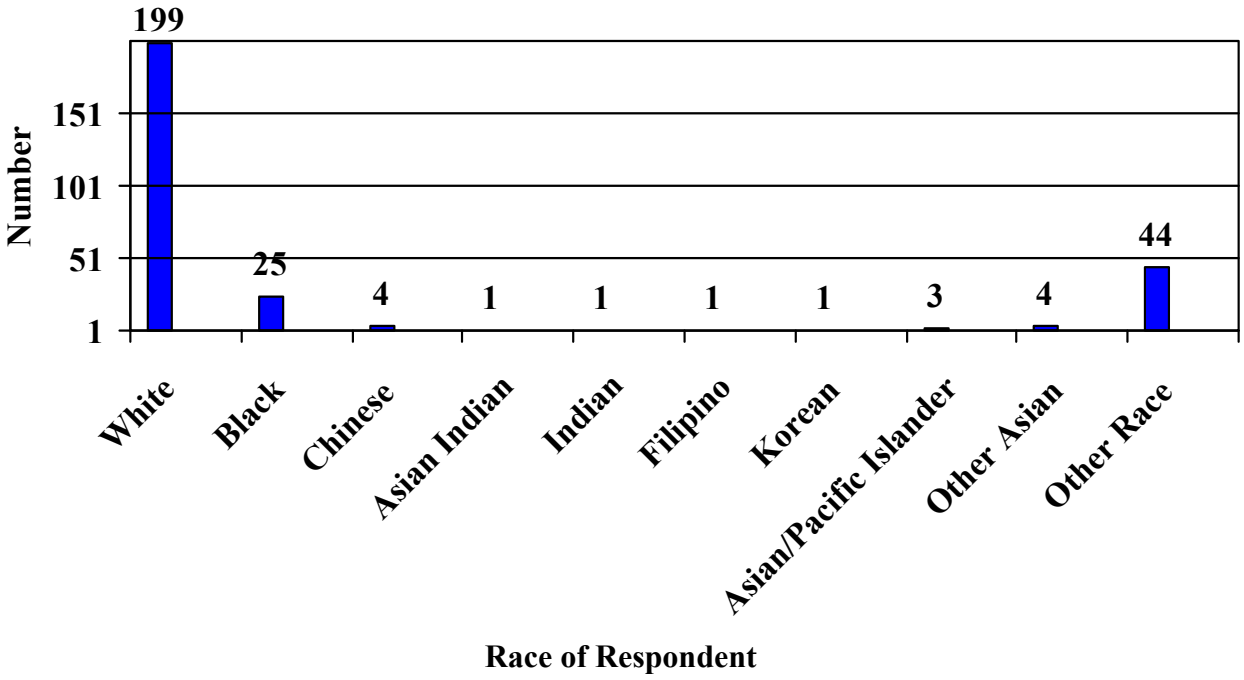
Age of respondent (Figure 26)



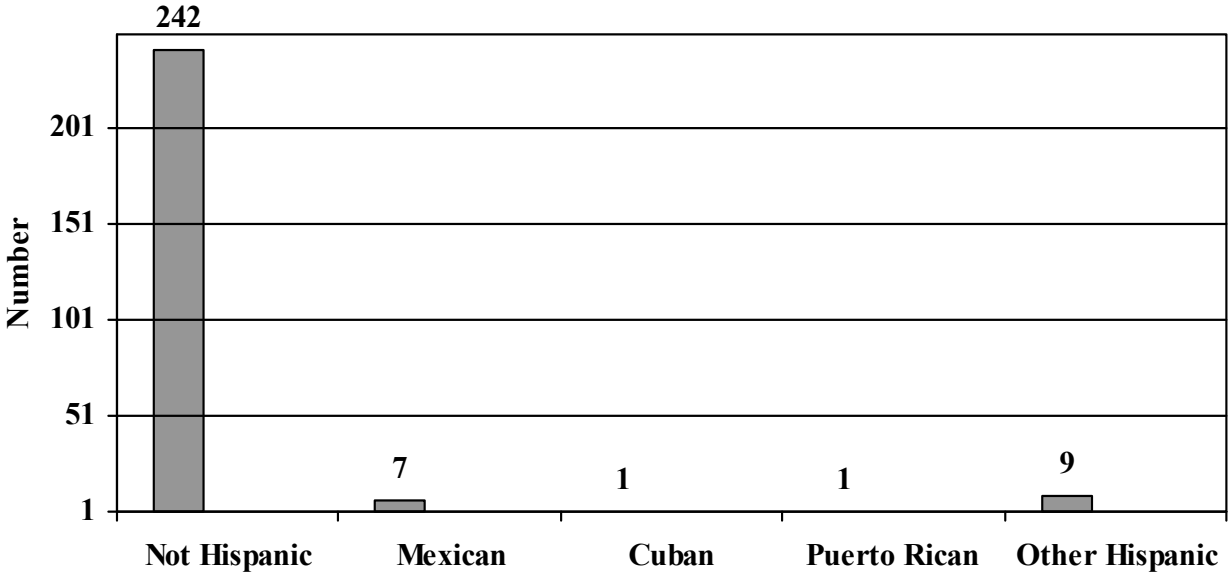
Sex of respondent (Figure 27)



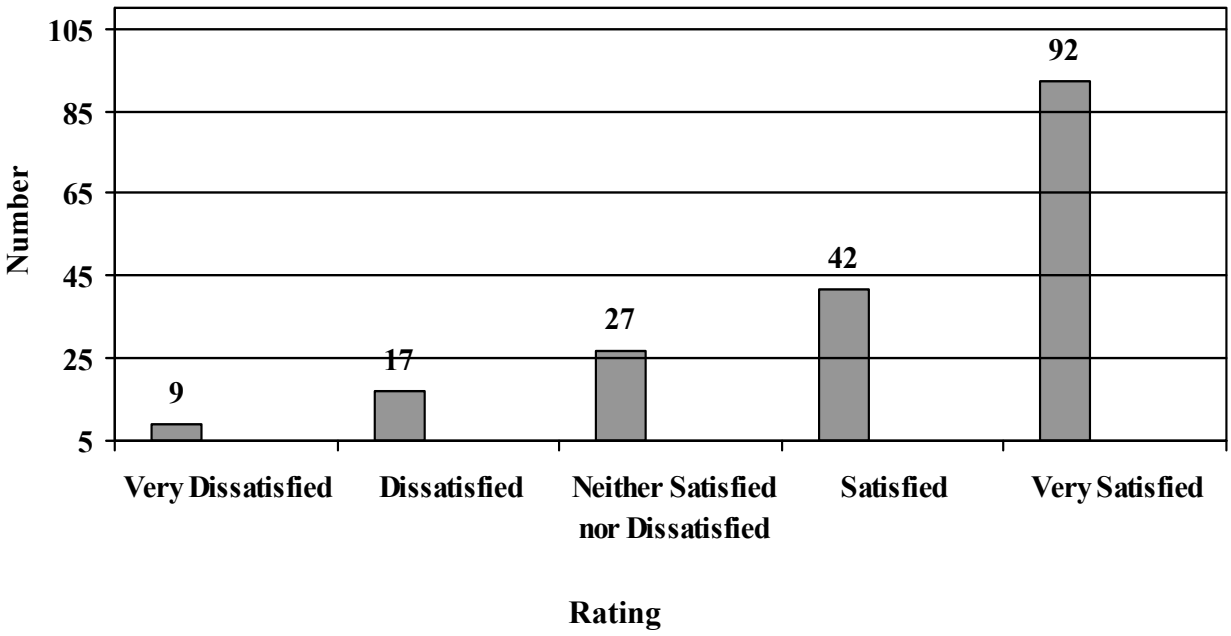
Race of respondent (Figure 28)



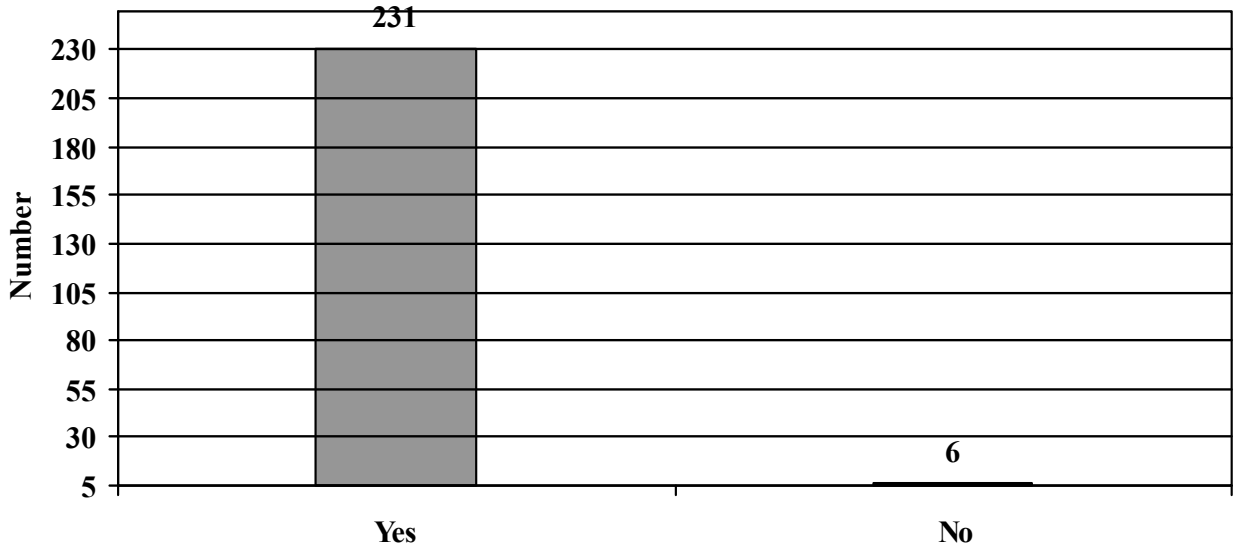
Ethnic background of respondent (Figure 29)



Respondent's rating of ASQ system (Figure 30)

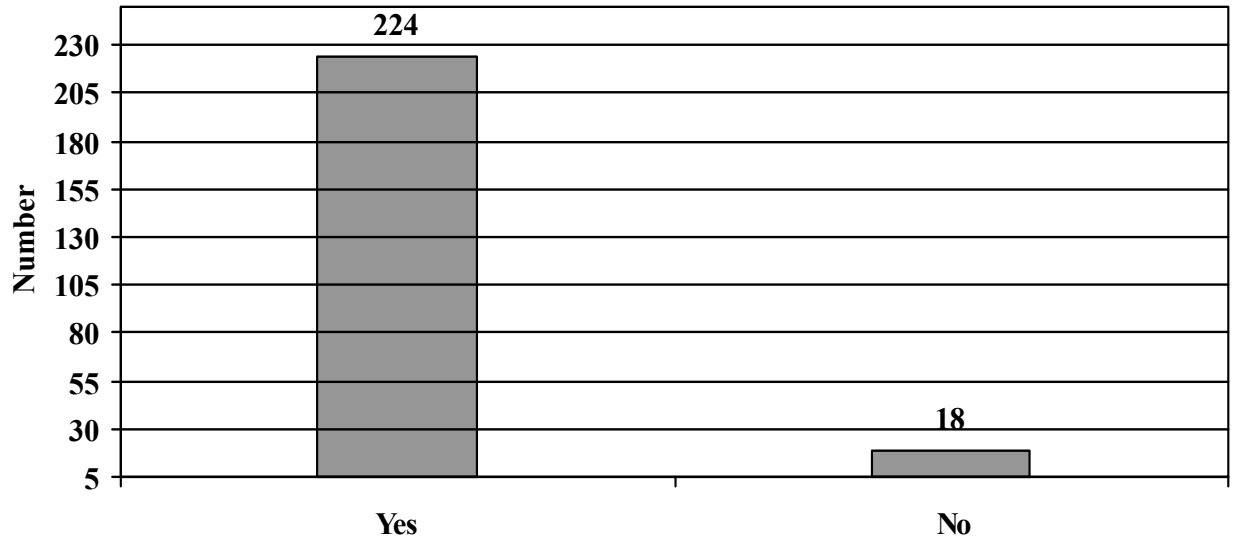


Respondent could understand the computer (Figure 31)



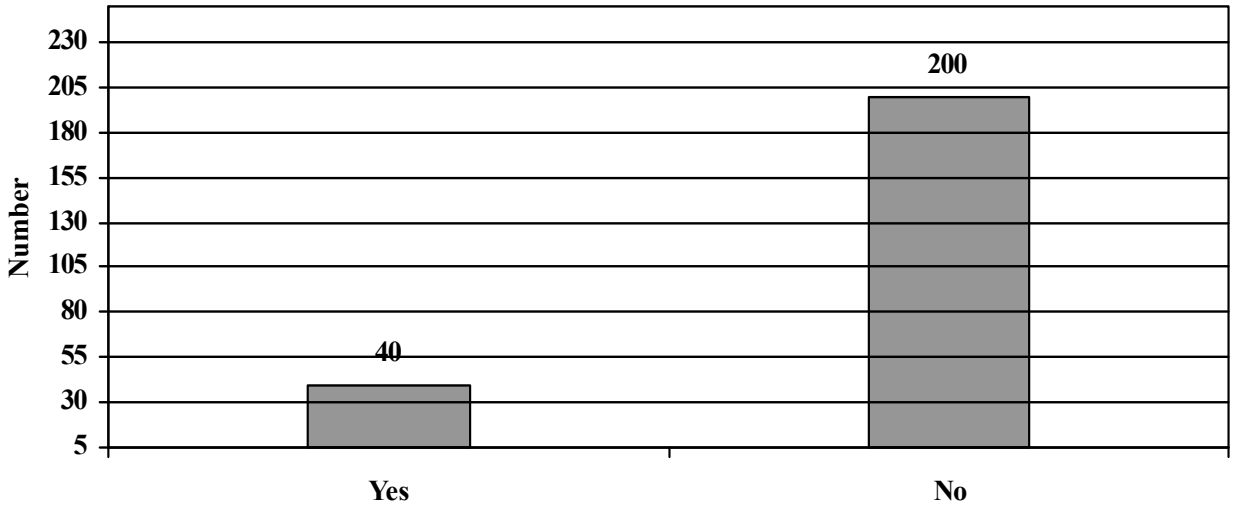
Respondent's Reported Ability to Understand the Computer

Respondent believed the computer could understand (Figure 32)



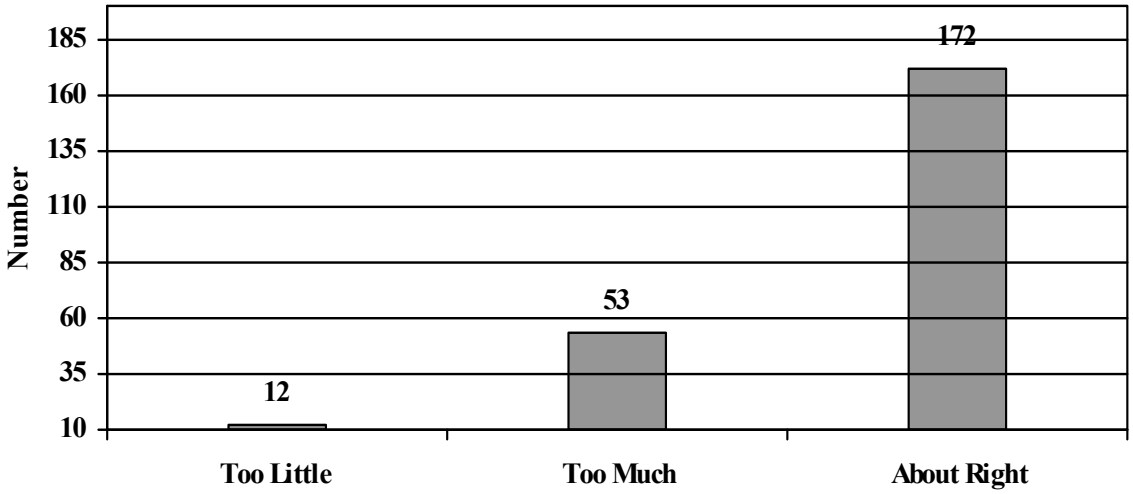
Respondent Reported the Computer Could Understand

Respondent was confused or frustrated (Figure 33)



Respondent Reported Being Confused or Frustrated

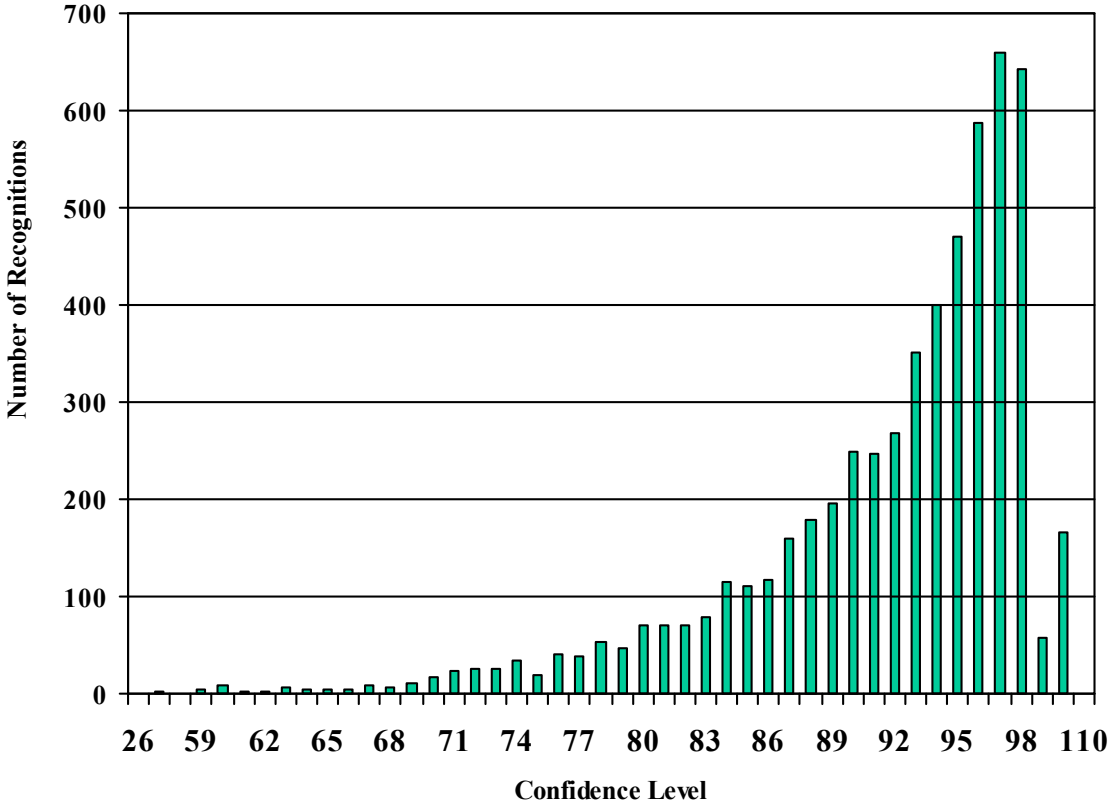
Respondent's belief about the amount of time afforded to answer questions (Figure 34)



Respondent Believed the Amount of Time Afforded to Answer Was

Figure 35 displays the confidence levels reported by the speech recognition software for all utterances in all calls.

Confidence level of Automatic Speech Recognizer (Figure 35)



To fully assess the issues related to the usability of the ASQ, we conducted regression analyses of the satisfaction and performance measures. The next section presents these results.

4.4.1 Regression Analyses of Performance Measures

This analysis was conducted in two steps. The first step involved conducting bivariate regressions of each performance measure on a single predictor variable. The performance measures used in these analyses include: (1) the overall ratings of the ASQ, (2) whether the ASQ was confusing and frustrating, (3) whether the system afforded right amount of time to answer, (4) the length of the calls in seconds and (5) the mean length of time in seconds that the respondent spent answering individual questions on the ASQ. Several other performance measures were also examined, including the respondent's ability to understand the ASQ and the respondent's perception of whether the ASQ understood them. However, the variance on these items was too low to do meaningful analysis (e.g., only six respondents reported they did not understand the ASQ).

The predictor variables that were included in this analysis include the performance measures described above, as well as a number of other demographic and call characteristics. Table 9 provides a complete list of variables included in the analysis.

Variables included as predictor variables in the regression analyses (Table 9)

Number in household	Mean confidence level
Length of call	Positive overall rating of ASQ
Mean time required by items	Respondent reported understanding computer
Number of retries for silence	Respondent found amount of time afforded to answer to be "just the right amount"
Number of retries for invalid response	Respondent found amount of time afforded to answer to be "too much time"
Number of retries for silence or invalid response	Respondent reported computer understood them
Respondent is female	Respondent found ASQ confusing
Age of respondent	Call occurred on a Monday
Respondent is white	Call occurred on a Wednesday
Respondent is black	Call occurred on a Thursday
Respondent is "other" race	Call occurred on a Friday
Household has at least two persons not same race	Call occurred on an afternoon
Respondent is not Hispanic	Call occurred on an evening
More than one Hispanic in the household	

Note–R stands for Respondent.

The second step was to conduct a multiple regression for each performance measure on those predictor variables that were statistically significant in the bivariate analysis. For each performance measure, regressions were estimated that started with all combinations of the basic demographic variables (household size; gender; race; Hispanic) that were significant in the bivariate analysis. Once the significant demographic predictors were determined, then characteristics of the interview were added, one at a time. Those variables that were statistically significant were kept in the equation until a final model was estimated which included only coefficients that were statistically significant.

An alternative procedure would be to estimate an equation with all of the variables and then eliminate non-significant variables, as done in the response rate analysis. This approach, however, becomes very cumbersome with the large number of variables used in the present analysis. Starting with the simplest model allows for more systematic exploration of different combinations of variables.

The bivariate analyses provide an overall picture on what is correlated with performance. This includes not only how the predictors are related to the performance measures, but also how these measures are interrelated. The multiple regression analysis provides information on the most important variables, once accounting for inter-item correlations.

Table 10 provide the results of the univariate regressions. In these tables, a minus sign indicates a negative relationship (i.e., as the value for the row variable changed, the value for the column variable tended to change in the opposite direction), a plus sign indicates a positive relationship, and a blank indicates no statistically significant relationship. An "x" simply indicates that the row variable and the column variable are the same.

Tables 11 - 15 provide the results of the multivariate regressions for each of the performance measures. For each performance measure, two models were estimated in which all predictor variables attained statistical significance. The two models differed in the combination of variables that were entered.

The one result that permeates across all of the results is that performance was affected by the number of people in the household. For example, Table 11 shows that respondents whose households had a large number of people tended to:

1. Give lower overall satisfaction ratings to the ASQ;

    - Rate the ASQ as confusing;
    - Dislike the amount of time afforded to answer the questions;
    - Have calls that took a relatively long time; and
    - Take an above-average amount of time to answer each survey item.

Perhaps, the questions on the census form seemed repetitive and wearisome to respondents who had to answer them repeatedly for many household members. These results fit well with the results reported by Cole and colleagues (1995) who found that respondents tended to dislike the ASQ data collection method when the data collection procedure required a relatively long time.

Related to this is that the respondents' level of satisfaction was influenced by the length of their calls. The univariate results (Table 11) show that respondents whose calls took a long time tended to:

- Give lower overall satisfaction ratings to the ASQ;
- Rate the ASQ as confusing; or
- Dislike the amount of time afforded to answer the questions.

The length of time to complete the questionnaire is highly related to the number of people who lived in the household. Thus, the multiple regression results do not include both length and number of people in the household in the same equation. This clearly shows, however, that household size is driving a lot of the problems users had by determining how long they had to stay on the phone, which in turn, affected other satisfaction measures (e.g., overall satisfaction – Table 12).

Summary of results of univariate regression analyses (Table 10)

	R rated ASQ positively overall	R rated ASQ as not confusing or frustrating	R found amount of time afforded by ASQ as appropriate	Length of call (sec)	Mean time for individual items (sec)
Number in household	-	-	-	+	+
Length of call	-	-	-	x	-
Mean time required by items				-	x
Number of retries for silence		-		+	
Number of retries for invalid response	-	-	-	+	
Number of retries for silence or invalid response	-	-		+	
Respondent is female	+				
Age of respondent	+		+	-	
Respondent is white		+			-
Respondent is black	+				
Respondent is “other” race		-	-		
Household has at least two persons not same race		-		+	
Respondent is not Hispanic					-
More than one Hispanic in the household		-		+	
Mean confidence level		+			
Positive overall rating of ASQ	x	+	+	-	
Respondent reported understanding computer					-
Respondent found amount of time afforded to answer to be "just the right amount”	+	+	x	-	+
Respondent found amount of time afforded to answer to be "too much time”					-
Respondent reported computer understood them			+		
Respondent found ASQ confusing	-	x	-	+	
Call occurred on a Monday				-	
Call occurred on a Wednesday				-	
Call occurred on a Thursday			-		
Call occurred on a Friday				+	
Call occurred on an afternoon					+
Call occurred on an evening					+

Note–R stands for Respondent.

Final results of linear regression analysis predicting how satisfied respondent was with the computerized questionnaire (scale of one to five) (Table 11)

	Model 1	Model 2
Intercept	3.20*** (.43)	3.46*** (.41)
Demographics		
Age	.01*** (.005)	.01*** (.005)
Gender		
Female	.34** (.15)	.37*** (.14)
Race		
Black	.56** (.25)	.48** (.22)
Number in household	-.18** (.09)	—
Call Characteristics		
Length of call	—	-.001** (.0003)
R found ASQ confusing	—	-.80*** (.23)
R^2	.10	.18

Note: See Table 9 for complete list of variables tested for inclusion in the regression model.

*p < .10
**p < .05
***p < .01

Final results of logistic regression analysis predicting that respondent described the questionnaire as not confusing (Table 12)

	Model 1	Model 2
Intercept	1.67*** (.47)	2.90*** (.36)
Demographics		
Race		
White	.87* (.49)	—
Other race	—	-1.71*** (.43)
More than one race in the household	-.94* (.54)	—
More than one Hispanic in the household	-.83* (.50)	-1.09** (.55)
Number in household	-.24* (.13)	-.32*** (.12)

Note: See Table 9 for complete list of variables tested for inclusion in the logistic regression model.

*p < .10
**p < .05
***p < .01

Final results of logistic regression analysis predicting that respondents said they had the right amount of time to answer the questions (Table 13)

	Model 1	Model 2
Intercept	-.72 (.73)	.62 (.70)
Demographics		
Age	.03** (.01)	.03** (.01)
Number in household	-.30* (.16)	-.29* (.15)
Call Characteristics		
Computer understood respondent	-1.30** (.54)	—
Respondent found ASQ confusing	—	-1.08** (.48)

Note: See Table 9 for complete list of variables tested for inclusion in the logistic regression model.

*p < .10
**p < .05

Final results of linear regression analysis predicting the length of the call (in seconds) (Table 14)

	Model 1	Model 2
Intercept	303.99*** (33.21)	429.19*** (53.29)
Demographics		
More than one race in the household	68.83* (35.88)	—
More than one Hispanic in the household	—	239.90*** (56.75)
Number in household	256.06*** (12.25)	233.22*** (15.48)
Call Characteristics		
Call on Friday	59.67* (34.03)	—
Mean length of time to complete each item	—	-9.54*** (2.00)
Number of retries for silence	13.10** (5.60)	18.43*** (6.80)
R^2	.85	.77

Note: See Table 9 for complete list of variables tested for inclusion in the regression model.

*p < .10
**p < .05
***p < .01

Final results of linear regression analysis predicting the mean length of time required to complete an item on the ASQ(in seconds) (Table 15)

	Model 1	Model 2
Intercept	24.79*** (2.04)	26.16*** (2.54)
Demographics		
Race		
White	-1.65** (.83)	-1.52* (.73)
Ethnicity		
Not Hispanic	-4.51** (1.89)	-3.79** (1.75)
Number in Household	-1.08*** (.21)	—
Call Characteristics		
Length of call	—	-.004*** (.001)
Call in the afternoon	1.09** (.54)	—
R^2	.20	.22

Note: See Table 9 for complete list of variables tested for inclusion in the regression model.

*p < .10
**p < .05
***p < .01

Paradoxically, respondents tended to spend a shorter amount of time on individual items when it took them longer to complete the instrument. This result suggests that the main reason that some calls were lengthy was that they included a large number of questions, not that the individual questions took a long time to answer. It may also indicate that respondents went faster with practice. The more times they were asked a particular question (e.g., age), the faster they were able to complete the item.

The bivariate results suggest that the respondents' opinions of the ASQ was associated with the number of times that the system had to repeat questions, either because the respondent failed to respond or because the speech recognition software returned a subthreshold confidence level for the recognition attempt. Respondents with higher numbers of retries for silence tended to rate the ASQ as confusing or frustrating. Perhaps their silence reflected their uncertainty about how to answer some of the questions. Respondents with higher numbers of retries because of invalid responses tended to be more generally dissatisfied with the ASQ, more likely to rate the ASQ as confusing, and more likely to find that the amount of time afforded to respond was inappropriate. These respondents may have been frustrated by the fact that the system repeated questions that they thought they had just answered, when the speech recognition software returned a subthreshold confidence level. Table 14 also shows that respondents with higher total number of retries, for both silence and for invalid responses, tended to give low overall ratings of the ASQ,

and to rate the ASQ as confusing. The table also shows that larger numbers of retries, whether for silence or invalid responses, tended to prolong the ASQ. Thus, retries tended to diminish the respondents' experience of the ASQ, perhaps by making the survey seem overly long, confusing, unnatural, or unlike human conversation.

The multivariate results indicate that re-tries for silence was a significant predictor of the length of the survey. The more re-tries the respondent had to enter, the longer the survey took (Table 15). The number of re-tries, however, was not statistically significant in the multiple regressions for overall satisfaction, whether the ASQ was confusing and whether the appropriate amount of time was given to the respondent. This suggests that many of the problems correlated with re-tries was related to increasing the length of the survey.

The respondents' Hispanic ethnicity affected their interactions with the ASQ. Hispanic respondents tended to spend more time per item than others (e.g., Table 16). Respondents from households with more than one Hispanic member tended to have relatively long calls and found the questionnaire more confusing. Thus, the ASQ was not totally "speaker independent." Apparently, the respondents' accents and ethnicity may be significant limitations in the performance of an English language ASQ. Another possibility might be issues these respondents may have with the Hispanic and Race questions. Many Hispanic respondents find the race question to be confusing, since it does not allow "Hispanic" as a racial category.

The respondents' other demographic characteristics also had some impact on their opinions of the ASQ. Female respondents tended to give the system higher overall satisfaction ratings. Older respondents tended to give the system higher overall satisfaction ratings and to find that the ASQ afforded them the appropriate amount of time to answer. White respondents tended not to find the ASQ confusing, and to spend less time answering the individual items in the ASQ. Black respondents tended to give the system higher overall satisfaction ratings. However, respondents who identified themselves with a race other than white or black (whether or not they also selected white or black) tended to find the ASQ to be confusing. Perhaps, these respondents had some trouble thinking through the question about race, or did not expect its "choose all that apply" format.

The racial complexity of the household also affected the interaction between the respondent and the system. Respondents from households in which there were people of different races tended to find the ASQ to be confusing and tended to have longer calls. This result may reflect the nature of the race questions on the census form. Respondents who could not answer the question the same way for everyone in the household may have had to devote effort to ensuring that they entered their data accurately.

The bivariate results show how the different performance measures were related to one another (Table 10). Respondents who found the ASQ to be confusing or frustrating tended to give lower overall satisfaction ratings to the system, to rate the amount of time that the system afforded them to answer as too much or too little, and to have longer calls. Respondents who gave high overall ratings to the ASQ tended also to rate the system as not confusing and as affording them the right amount of time to answer the questions. Their calls tended to be short. Again, these results fit well with those of Cole and colleagues (1995), who found that respondents tended to be most satisfied with speedy data collection.

The results of RMIE can guide future attempts to field IVR-based surveys. The results suggest that these surveys should be designed to be efficient, particularly for respondents whose households have characteristics that might prolong the data collection time. In RMIE, the ASQ protocol closely matched the format of the printed census short form. Designers of future versions of a census short form speech application might strive to streamline the data collection process so that respondents can provide their data in the least amount of time. The format that works best for paper forms may not be the most efficient format for an IVR-based survey. For example, future ASQ protocols could include questions such as, "Are all members of the household also Mexican?" so that respondents did not have to report the ethnic background of every member of the household separately when doing so would be repetitious. Respondents might also be required to enter a short alias for the 22-digit census identification number.

The results also suggest that the ASQ should better accommodate Hispanic respondents. Perhaps the system should ask callers early in the interview whether they would prefer to be interviewed in Spanish. Those who answer affirmatively could be transferred to a bilingual CATI operator or a Spanish ASQ. In this way, Hispanic respondents who prefer to speak Spanish could do so, even if they are capable of conversing in English as a second language.

Speech recognition technology is rapidly improving. Interactions between users and IVR systems will eventually resemble natural conversation. The current project used the Antares model speech processing board from Dialogic, Inc., an Intel subsidiary. (http://www.intel.com/network/csp/trans/dialogic.htm) Since this ASQ was designed, Dialogic has marketed new digital speech processor products, and several companies have marketed improved speech recognition software packages. More efficient speech processing algorithms and faster computing are likely to lead to better speech applications.

In a separate study, the firm Speechworks (2000) analyzed the voice files containing the respondents' recorded names and their spelling. Speechworks' report from that study suggested that the ASQ could be improved in a number of ways. The software routines for endpointing could be improved to eliminate any noise at the end of the utterances. Barge-in could be enabled at some points in the ASQ, allowing respondents to interrupt questions. Perhaps some questions could be reworded to encourage respondents to reply with words that are in the expected vocabulary. Replies that received subthreshold confidence scores from the speech recognizer might trigger followup questions such as, "I thought you said 'daughter.' Is that correct?" rather than simply a repetition of the question. Respondents could then confirm their responses with "yes" or "no," rather than repeat the word that received the subthreshold confidence score.

In the next few years, versions of Extensible Markup Language (XML) for voice applications (called Voice XML) will lead to web-based voice applications and new development and data management environments for IVR systems. Quite possibly, IVR-based data collection methods will be used routinely in the near future for many large-scale data collection efforts like the decennial census.

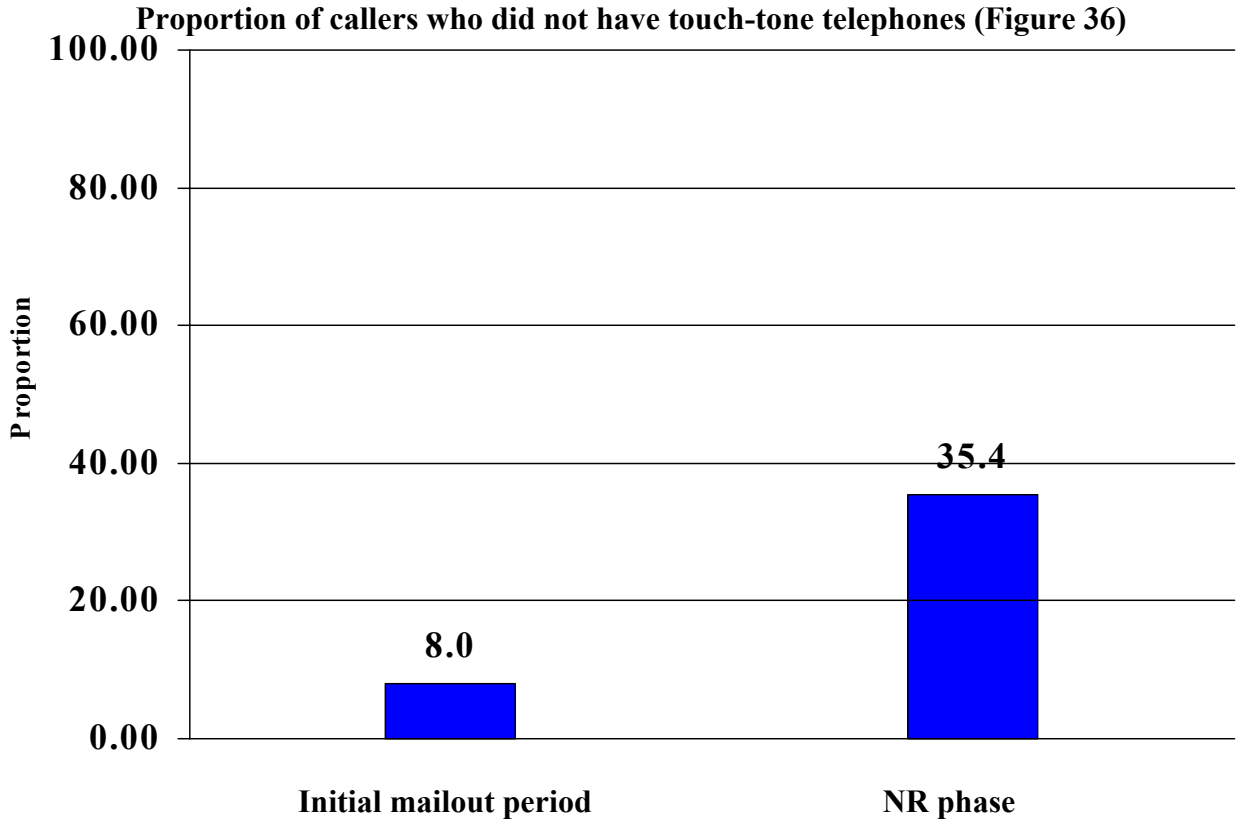
4.4.2 Open-ended questions

Appendix F displays the transcriptions of the replies to the open-ended user satisfaction questions. The statements suggest that some respondents thought that the ASQ required too much time. Another frequent complaint was that the ASQ did not provide a way to correct response immediately after they were given. Instead, respondents had to wait until the end of the ASQ to correct any errors. Several respondents had complaints about the decennial census itself and the nature of the questions on the short form.

Some respondents liked the ASQ because they believed that it was a speedy and convenient alternative to the paper forms. No respondent mentioned being unable to fill out the paper forms because of limited literacy skills or disability. The issue of confidentiality was not mentioned. The results suggest that respondents seek a way to dispense with their census duty expeditiously and conveniently.

4.4.3 Touch-tone telephones

Figure 36 shows the proportion of callers to the ASQ who had no touch-tone telephone. The proportion was much higher during the NR phase, as compared with the initial mailout period. Callers without touch tone telephones had to enter their 22-digit census identification number and telephone number verbally. In both periods only 1.8 percent of the respondents who had no touch-tone telephone were able to complete the interview without being transferred to the CATI operator.



5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Recommendations: Offering an Alternative Response Mode

The results of the RMIE can help guide future use of computer-mediated response modes and incentives in the decennial census. The results address the questions:

- Can offering alternate, computer-mediated response modes increase overall response rates?
- Do respondents using alternate, computer-mediated response modes tend to provide good quality data?
- Are the costs involved in offering alternative response modes commensurate with any advantages they offer?

Overall response rates did increase when respondents were offered the CATI and Internet alternative modes, as compared with the control group. The increase in overall response rates was small and occurred only when the respondents were not offered an incentive. When an incentive was offered, overall response rates went down slightly, to about the same level as that of the control group. These alternative response modes also seemed to reduce the amount of missing data for particular items; that is, the item nonresponse rates tended to be higher for mail questionnaires as compared with CATI and Internet questionnaires.

Considering the scope of the experiment, the major drawback to CATI is its cost. CATI involves a number of expenses that the other modes do not require, such as the costs associated with the interviewers, CATI equipment and software, and the 800 telephone line. The interviewer costs are increased by the time that they must spend unoccupied, waiting for calls to come in. Early in the calling period, calls come in frequently; however, toward the end of the calling period, the calls become less frequent and the interviewers spend increasing amounts of time waiting. In RMIE, interviewers had to be given other tasks while they were waiting for calls. They transcribed voice files from the ASQ and placed callbacks to CATI respondents whose questionnaires were not completed. Nonetheless, the costs of employing telephone interviewers several weeks after the mailing can be quite high.

However, CATI may provide some cost savings within the context of a large-scale census data collection effort. CATI data collection saves the costs for return postage and data capture associated with mail surveys. Also, CATI did seem to improve some aspects of data quality and exhibit less missing data than the mail survey on certain items.

It is difficult to assess these trade-offs precisely. Based on Westat's experiences with the relative costs of CATI, in-person and mail surveys, and this experiment, we suspect that CATI poses a significant increase in cost relative to the current census procedures, unless the costs were offset by a large increase in the response rate. The RMIE results suggest that offering a CATI response

mode alternative does not bring about such a large increase in the response rate. However, in a Decennial Census, CATI might provide cost efficiencies as a component in a telephone questionnaire assistance program.

Like CATI, the Internet mode yielded relatively high data quality. There was also a relatively low rate of missing data on key items. When an incentive and insert were not included, the response rate was approximately one to two percentage points higher than that of the CCG. Relative to the census mail procedure, the costs of fielding a web survey are likely to be relatively modest. The primary additional cost associated with the Internet, relative to mail, involves the development and maintenance of the software and hardware. However, this cost is fixed and does not increase as more data are collected. Web surveys also have lower postage and processing costs than mail surveys do. Data quality could be improved further with the introduction of automated edits.

It is difficult to know where the "break-even" point might be (i.e., how much data need to be collected to cover the development of the web survey). Based upon conservative assumptions and the data from RMIE, one might save between one to five million dollars in postage costs alone if between 3 percent and 15 percent of the sample uses the web rather than the mail survey. This estimate assumes that the postage to mail back the short form is 30 cents and 110 million households must be enumerated (3 percent x 110 million households x .30 cents postage = $990000; 15% x 110 million households x .30 cents = $4.95 million). This savings would more than offset the costs required to design, develop and maintain the web survey. Of course, the web survey would also produce savings related to reduced processing (receipt and scanning). Given this crude calculation, we would expect that the Internet would be cost-effective even if a relatively small proportion of respondents used it. The Census Bureau may be required to offer an Internet response option if for no other reason than public perception. Also, offering a web survey would offer additional cost savings if it increased the overall response rate, as it did in RMIE. Fewer followup field interviews would be required.

The implications of this experiment for the use of the ASQ are complex. Data quality was the lowest for this response mode, both in terms of response rate and missing data items. Much of these missing data were due to individuals hanging up relatively early during the interview. With respect to costs, the ASQ has fixed costs related to purchasing the hardware, developing the software and maintaining the data collection site. There are other costs if operator assistance is provided for those individuals who cannot complete the questionnaire using ASQ. There are also additional data-processing costs because of the need to transcribe information that the speech recognizer could not code. Therefore, an ASQ is more costly than an Internet survey. It is unclear how ASQ costs compare to those of CATI or mail questionnaires.

With respect to data quality, this study points to at least two concerns with the ASQ for the use on future censuses. One is the packaging of the offer to use this mode. Many of the issues discussed for the Internet also apply for the ASQ (i.e., use of an insert, types of messages to promote use of the ASQ). An additional issue is whether (and how) to inform respondents that they would be providing their data to a computer. The RMIE mailings did not notify ASQ households that the telephone number was for an ASQ. Some of the negative reaction to the ASQ may have been avoided if respondents made the call with the expectation that they would be interacting with an automated system. Future work should consider alternative ways to present this information to respondents. If respondents understand they will be providing data to a computer before calling, they may react more positively when they encounter the ASQ system.

Another concern revolves around the design of the ASQ interview. Several tasks were difficult to complete or took more time than desired on the ASQ. This likely affected the quality of the data with this mode. Issues that may have led to problems include: (1) entering a 22-digit ID, (2) reporting and spelling out the names of all persons in the household and (3) reporting race using information printed on the paper questionnaire.

Some of these issues were a function of the special nature of this experiment within Census 2000. For example, shortening the ID may be possible if a crosswalk could be developed between the full 22-digit census number and a shorter number that would be easier to enter. Also, the ASQ may become easier to use as the technology of speech recognition becomes more sophisticated. For example, the ASQ did not rely on recognizing the responses to every question. The responses to the questions on race and certain other topics were recorded and later transcribed. Improved capabilities to recognize speech, especially words embedded within a sentence (e.g., reports of multiple races), would allow for easier interaction between the respondent and the computer.

5.2 Recommendations: Offering an Incentive With an Alternative Response Mode

The RMIE results show that the inclusion of a calling card with an insert was extremely effective in promoting the use of the alternative response mode. Comparisons between the incentive and no-incentive conditions reveal that the incentive was associated with three to four-fold increases in the rate of using the alternative mode.

At least some portion of this effect is probably attributable to the insert, which drew the respondents' attention to the availability of the alternative mode. The non-incentive condition relied solely on the census cover letter to inform respondents about the availability of the computer-mediated mode. Many respondents in the no-incentive panels probably did not read the letter. The insert, by contrast, prominently called the respondents' attention to the computer-mediated alternative mode. The insert and calling card may account for some of the effects observed in the incentive condition.

However, this increase seemed to come at some cost to the overall response rate with one to two percent *fewer* people responding when an incentive was offered. In both the CATI and Internet conditions, the overall response rates, once factoring in the mail responses, were lower in the incentive panels than in the no-incentive panels. This reduction may be due to the fact that the calling card incentive makes the response task more complicated. If the alternative modes are not available at the time the respondent tries to use them, the respondent may not follow up in all cases to complete the questionnaire at a later time. The one advantage of a mail questionnaire is that it can be filled out the moment the package arrives. Completing a CATI questionnaire requires the use of a telephone and the availability of a CATI operator. A web survey requires access to a computer that has Internet access. If these are not available at the time the respondent attempts to fill out the questionnaire, then some persons may simply never respond.

This result may also be indicative of a relatively weak effect of the calling card as an incentive. In fact, many respondents whose calling cards were activated never used them, suggesting that the calling card may not have been a universally powerful incentive.

Given the success of the insert and incentive to promote the use of an alternative mode to respond to the census, this option should be considered in future research. This research should carefully consider both the role the insert and incentive separately play in the respondent's decision to participate. It would be useful to better understand the relative effects of the calling card (incentive) and the insert on the respondents' decision to use the alternative mode. The use of just an insert, without any incentive, has a number of economical and logistical advantages for the census. Research is needed into the best ways to present the alternative through either the letter or an insert. Related research would test different types of messages given to the respondent.

6. REFERENCES

Baker, R. P., Bradburn, N. M., and Johnson, R. A. (1995) Computer-assisted personal interviewing: An experimental evaluation of data quality and cost. *Journal of Official Statistics, 11,* 413-431.

Benjamini, Y. and Y. Hochberg. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *Series B,* 57, 289-300.

Blyth, B. (1997) Developing a speech recognition application for survey research. In: Lyberg, L., Biemer, P., collins, M., deLeeuw, E., Dippo, C., Schwarz, N., and Trewin, D. *Survey measurement and process quality.* New York: Wiley, pp.249-266.

Bouffard, J. & Guarino, J. (2001) *Results from the Internet Usage Survey.* Unpublished manuscript.

Bureau of the Census (1999, August) *Census 2000 dress rehearsal evaluation summary.* Planning, Research, and Evaluation Division.

Bureau of the Census (2000) *Memorandum Series #K-3*. From Michael J. Longini, Chief, Decennial Systems and Contracts Management Office. Subject: Definition of a blank form by DCS 2000.

Cole, R. A., Novick, D. G., Vermeulen, P. J. E., Sutton, S., Fanty, M., Wessels, L. F. A., deVilliers, J.H., Schalkwyk, J., Hansen, B., and Burnett, D. (1995) Experiments with a spoken dialogue system for taking the U.S. Census. In: *Proceedings of the International Journal on Human Computer Interaction*, 1-39.

Couper, M. P., Singer, E., and Kulka, R. A. (1998) Participation in the 1990 decennial census: Politics, privacy, pressures. *American Politics Quarterly, 26,* 59-80.

Guarino, J. *Assessing the impact of differential incentives and alternate data collection modes on census response.* Unpublished manuscript.

Kirsch, I., Jungeblut, A., Jenkins, L. and Kolstad, S. (1993) *Adult literacy in America: A first look at the results of the National Adult Literacy Survey*. Washington DC: U.S. Department of Education, National Center for Educational Statistics.

National Telecommunications and Information Administration (2000) *Falling through the net: Toward digital inclusion.* Available at www.ntia.doc.gov.

Singer, E., Hippler, H.J. and Schwarz, N. (1992) Confidentiality assurances on surveys: Reassurance or threat? *International Journal of Public Opinion Research, 4*, 256-268.

Singer, E., Mathiowetz, N. A., and Couper, M. P. (1993) The impact of privacy and confidentiality concerns on survey participation: The case of the 1990 U.S. Census. *Public Opinion Quarterly*, 57, 465-482.

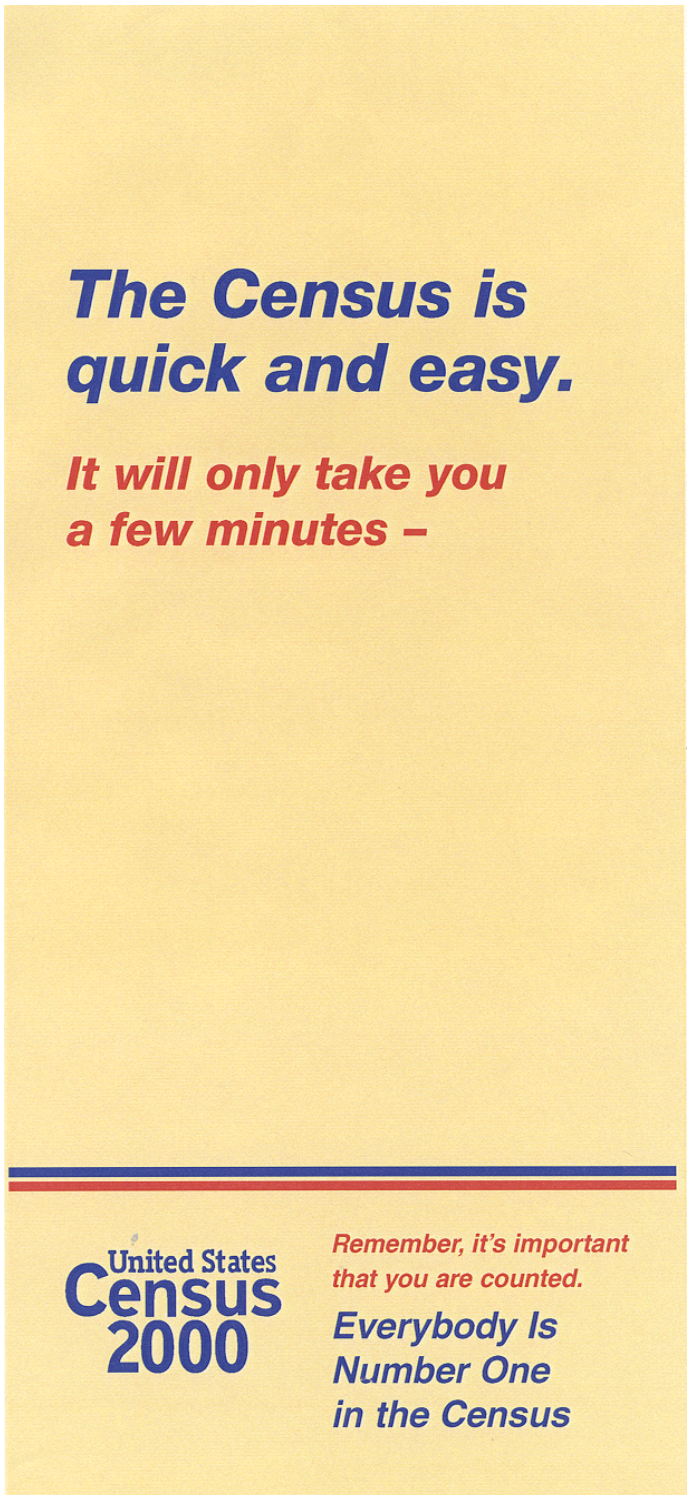
Speechworks (2000, September) *Census 2000 name recognition.* Unpublished manuscript.

Turner, C. F., Ku, L., Rogers, S. M., Lindberg, L. D., Plaeck, J. H., and Sonenstein, F. L. (1998) Adolescent sexual behavior, drug use and violence: Increased reporting with computer survey technology. *Science*, *280*, 867-873.

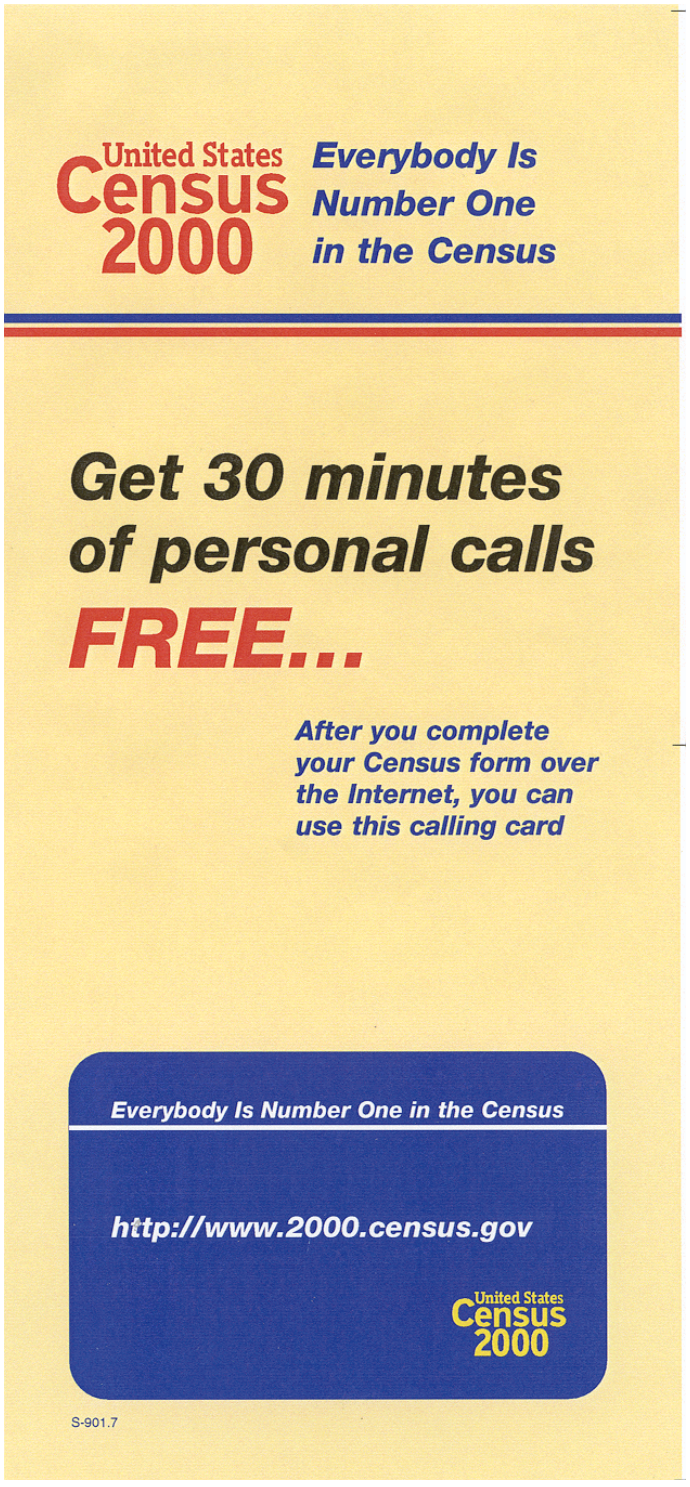
Turner, C. F., Miller, H. G., Smith, T. K. Cooley, P. C. and Rogers, S. M. (1996) Telephone audio-Casi: Preliminary evaluation of a new technology for surveys of sexual and other sensitive behaviors. *Technical Papers on Health and Behavior Measurement. Technical Paper Number 21.* Rockville, MD: Research Triangle Institute.

Appendix A

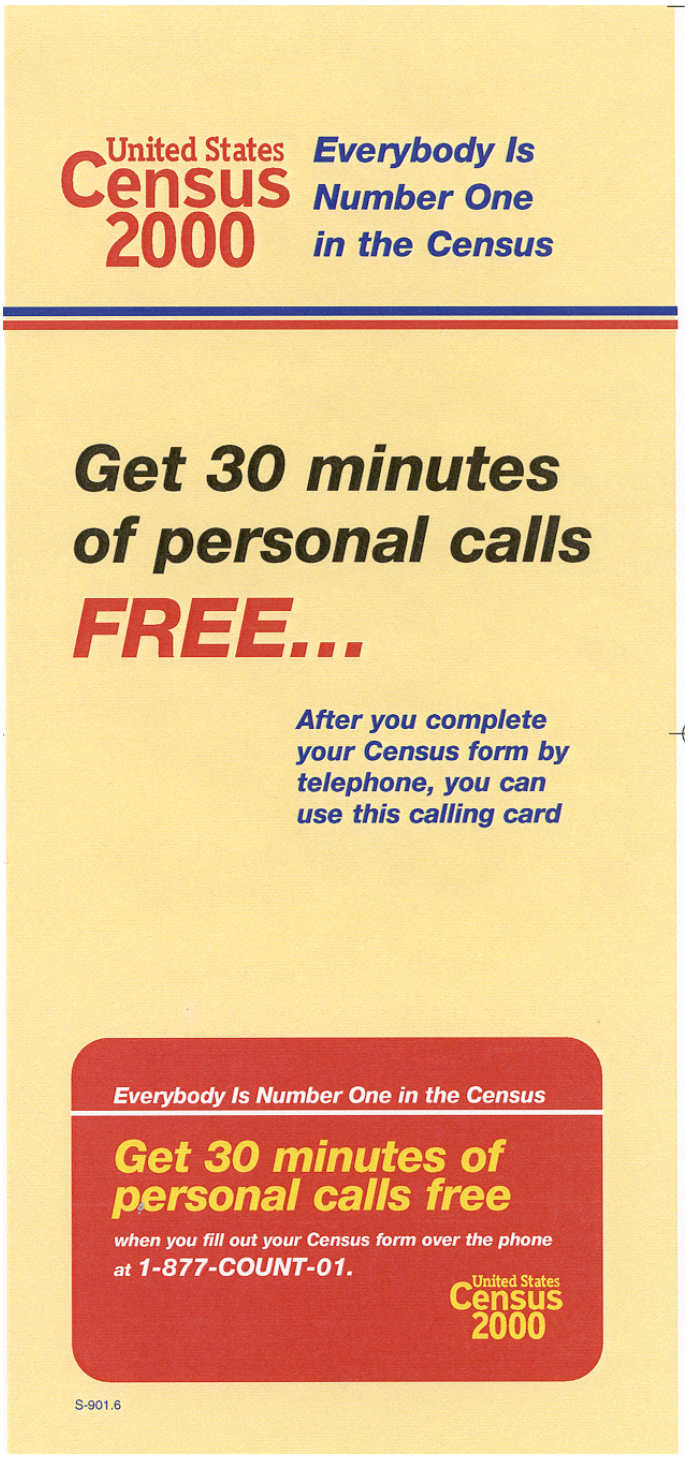
Inserts Included with Mailings



APPENDIX A


United States Census 2000
Everybody Is Number One in the Census
Get 30 minutes of personal calls FREE...
After you complete your Census form over the Internet, you can use this calling card
Everybody Is Number One in the Census
http://www.2000.census.gov
United States Census 2000
S-901.7

APPENDIX A


United States
Census
2000
Everybody Is
Number One
in the Census
Get 30 minutes
of personal calls
FREE...
After you complete
your Census form by
telephone, you can
use this calling card
Everybody Is Number One in the Census
Get 30 minutes of
personal calls free
when you fill out your Census form over the phone
at 1-877-COUNT-01.
United States
Census
2000
S-901.6

APPENDIX A


United States Census 2000
Everybody Is Number One in the Census
Get 30 minutes of personal calls FREE...
After you complete your Census form by telephone, you can use this calling card
Everybody Is Number One in the Census
Get 30 minutes of personal calls free
when you fill out your Census form over the phone at 1-877-53-COUNT.
United States Census 2000
S-901.5

Appendix B

ASQ Protocol

Final ASQ 2000 short form script

Test line: 1-877-286-3119

Revised March 7, 2000

Notes:

Panel 2 = ASQ Control Panel
Panel 5 = ASQ with calling card incentive
Panel 8A= ASQ with calling card, NRFU
Panel 8B = ASQ, no calling card, NRFU

All responses must be recorded for playback and verification and transcribed if necessary.

Feedback to the respondent is done by the recorded audio clips.

Some responses do not have to be recognized in real time. Spelled and spoken names will be processed by SpeechWorks in post-processing and ship the results back to Westat. These entries are noted by:

<record, transcribe, recognize (SpeechWorks)>

Other responses may be processed in batch mode at the end of the project to obtain information about the recognition confidence level needed for the ASQ usability analysis. These entries are noted by:

<record, transcribe, batch recognition later>

Open-ended responses from the satisfaction survey are marked with this entry:
<record, transcribe>

##
[chime] You have reached the Census Bureau's Computerized Questionnaire.

[testing] The data you provide is not confidential and will be used for software development.
(*March 13, 2000 and later*) Your answers are protected by law.

(*April 1, 2000 and earlier*)
You will be asked to provide information about yourself and persons living in your household on April 1, 2000, including:

(*April 2, 2000 and later*)
You will be asked to provide information about yourself and persons who were living in your household on April 1, 2000, including:

* **last name, first name and middle initial;**
* **sex**
* **date of birth**
* **age on April 1, 2000**
* **origin**
* **race**
* **relationship**

[chime] We will record your information. When you hear this beep <tone> please speak and keep your answers brief. Please keep the form with your questionnaire ID at hand to assist you with some of the questions. We will now begin.

Do you have a telephone with number buttons, yes or no? <tone>

if silence, then
Do you have a telephone with number buttons? Please say yes or no AFTER the beep. <tone>
if silence, then
transfer to CATI

Your questionnaire ID number is located above your address on the form mailed to you.

<If yes, then>
<all touchtone digits scenario>
buttons = true
ID22:

Please enter all 22 digits of your questionnaire id with the pushbutton keys on your telephone after you hear the beep.<tone2>

<accept input>
You entered $$$$$ dash $$$$$$$ dash $$ dash $$$ dash $$$ dash $$, Is this correct, yes or no? <tone>

If yes, then
goto VERIFY 1

If no, then
goto AGAIN

if silence, then
You entered $$$$$ dash $$$$$$$ dash $$ dash $$$ dash $$$ dash $$.
Is this correct? Please say yes or no after the beep. <tone>

If yes, then
goto VERIFY 1
If no OR silence, then
goto AGAIN
<end all touchtone digits scenario>

<begin punctuated touchtone digits scenario>
ID5: Please enter the first five digits of your Questionnaire ID with the pushbutton keys on your telephone after you hear the beep.<tone2>

<accept input>
You entered xxxxx. Is this correct, yes or no? <tone>

if silence, then
You entered xxxxx. Is this correct? Please say yes or no after the beep <tone>

If no, then
go to ID5

ID7: Please enter the next seven digits of your questionnaire ID with the pushbutton keys on your telephone after you hear the beep. <tone2>

<accept input>
You entered xxxxxxx. Is this correct, yes or no? <tone>

if silence, then
You entered xxxxxxx. Is this correct? Please say yes or no after the beep <tone>

If no then go to ID7

ID2: Please enter the next 2 digits of your questionnaire ID with the pushbutton keys on your telephone after you hear the beep. <tone2>

<accept input>
You entered xx. Is this correct, yes or no? <tone>

if silence, then
You entered xx. Is this correct? Please say yes or no after the beep <tone>

If no then go to ID2

ID3: Please enter the next 3 digits of your questionnaire ID with the pushbutton keys on your telephone after you hear the beep. <tone2>

<accept input>
You entered xxx. Is this correct, yes or no? <tone>

if silence, then
You entered xxx. Is this correct? Please say yes or no after the beep <tone>

If no then go to ID3

ID32: Please enter the next 3 digits of your questionnaire ID with the pushbutton keys on your telephone after you hear the beep. <tone2>

<accept input>
You entered xxx. Is this correct, yes or no? <tone>

if silence, then

You entered xxx. Is this correct? Please say yes or no after the beep <tone>

If no then go to ID32

ID222: Please enter the last two digits of your questionnaire ID with the pushbutton keys on your telephone after you hear the beep. <tone2>

<accept input>
You entered xx. Is this correct, yes or no? <tone>

if silence, then
You entered xx. Is this correct? Please say yes or no after the beep <tone>

If no, then go to ID222
goto VERIFY 2
<end punctuated touchtone digits scenario>

<If no, then>
<all spoken digits scenario>
IDV22:
Please say all 22 digits of your questionnaire ID without pausing after you hear the beep.<tone>

<accept input>
You said $$$$$ dash $$$$$$$ dash $$ dash $$$ dash $$$ dash $$, Is this correct, yes or no? <tone>

If yes, then
goto VERIFY 1

If no, then
goto AGAIN

if silence, then
You entered $$$$$ dash $$$$$$$ dash $$ dash $$$ dash $$$ dash $$.
Is this correct? Please say yes or no after the beep. <tone>

If yes, then

goto VERIFY 1
If no OR silence, then
goto AGAIN
<end all spoken digits scenario>

<punctuated spoken digits scenario>
IDV5: Please say the first five digits of your Questionnaire ID after you hear the beep. <tone>

<accept input>
You said xxxxx. Is this correct, yes or no? <tone>

if silence, then
You entered xxxxx. Is this correct? Please say yes or no after the beep <tone>

If no, then
go to IDV5

IDV7: Please say the next seven digits of your questionnaire ID after you hear the beep. <tone>

<accept input>
You said xxxxxxx. Is this correct, yes or no? <tone>

if silence, then
You entered xxxxxxx. Is this correct? Please say yes or no after the beep <tone>

If no then go to IDV7

IDV2: Please say the next 2 digits of your questionnaire ID after you hear the beep. <tone>

<accept input>
You said xx. Is this correct, yes or no? <tone>

if silence, then
You entered xx. Is this correct? Please say yes or no after the beep <tone>

If no then go to IDV2

IDV3: Please say the next 3 digits of your questionnaire ID after you hear the beep. <tone>

<accept input>
You said xxx. Is this correct, yes or no? <tone>

if silence, then
You entered xxx. Is this correct? Please say yes or no after the beep <tone>

If no then go to IDV3

IDV32: Please say the next 3 digits of your questionnaire ID after you hear the beep. <tone>

<accept input>
You said xxx. Is this correct, yes or no? <tone>

if silence, then
You entered xxx. Is this correct? Please say yes or no after the beep <tone>

If no then go to IDV32

IDV222: Please say the last two digits of your questionnaire ID after you hear the beep. <tone>

<accept input>
You said xx. Is this correct, yes or no? <tone>

if silence, then
You entered xx. Is this correct? Please say yes or no after the beep <tone>

If no then go to IDV222
goto VERIFY 2
<end punctuated spoken digits scenario>

:VERIFY 1

<verify ID>

<If no match, then>
The number you entered is not in our records.
:AGAIN
if buttons, then
go to ID5
else
go to IDV5

:VERIFY 2
<verify ID>

<if ID used before, then
if complete, then
We see from our records that you already provided your Census information. We are transferring you to an operator who will answer your questions.
else
We see from our records that you entered some information into this system. We are transferring you now to an operator who will take your information.
<transfer to OA>
>

<If no match, then>
See bailout specification

#####HOME_OWNER
[chime] We will now ask you about this property.

Is this property owned by you or someone in this household free and clear, without mortgage, yes or no? <tone>

if silence, then
Is this property owned by you or someone in this household free and clear without a mortgage? Please say yes or no after the beep <tone>

<if no, then>
Is this property owned by you or someone in this household with a mortgage or loan, yes or no? <tone>

if silence, then

Is this property owned by you or someone in this household with a mortgage or loan? Please say yes or no after the beep. <tone>

<if no, then>

Is this property rented for cash, yes or no? <tone>

if silence, then

Is this property rented for cash? Please say yes or no after the beep. <tone>

<if no, then>

Is this property occupied without payment of cash rent, yes or no? <tone>

if silence, then

Is this property occupied without payment of cash rent? Please say yes or no after the beep. <tone>

#######NAME & TELEPHONE

We will need your name and telephone number in case we need to contact you to understand or clarify an answer. Please say your first name after the beep.<tone>

<record, transcribe, recognize (SpeechWorks)>

if silence, then

Please say your first name AFTER you hear the beep. <tone>

<record, transcribe, recognize (SpeechWorks)>

Please say your last name after the beep. <tone>

<record, transcribe, recognize (SpeechWorks)>

if silence, then

Please say your last name AFTER you hear the beep. <tone>

<record, transcribe, recognize (SpeechWorks)>

Ok ... now

short = false

silence1 = false

Phone:

If buttons, then
if not short, then
Please enter your phone number, area code first, with the number buttons on your telephone now. <tone2>
else
if short or silence1, then
Please enter all ten digits of your phone number AFTER you hear the beep.<tone2>
else
if not short, then
Please say your phone number, area code first, by speaking one digit at a time now. <tone>
else
if short or silence1, then
Please say all ten digits in your phone number without pausing AFTER you hear the beep.<tone>

We have xxx <pause> xxx <pause> xxxx as your telephone number. Is this correct, yes or no? <tone>

telephone = false
if silence, then
silence1 = true
goto Phone

If no, then
If count(digits) < 10, then short = true
go to Phone

If yes, then telephone = true

NAME_LIST

(April 1, 2000 and earlier)
Next, you will be asked to list any other persons living at this address on April 1, 2000.
(April 2, 2000 and later)
Next, you will be asked to list any other persons who lived at this address on April 1, 2000.

Certain persons will be counted at other places, so DO NOT INCLUDE anyone who is:

away at college,
OR in a correctional facility, nursing home, or mental hospital on April 1, 2000,
OR in the Armed Forces and living somewhere else,
OR staying at another place most of the time.

In addition to yourself, are there any other household members that need to be counted, yes or no? <tone>

if silence, then

In addition to yourself, are there any other household members that need to be counted? Please say yes or no after the beep. <tone>

If yes, then

roster = 2
Please say the first name of person 2 after the beep. <tone>
<record, transcribe, recognize (SpeechWorks)>

if silence, then

Please say the first name of person 2 AFTER you hear the beep. <tone>
<record, transcribe, recognize (SpeechWorks)>

Please say the last name of person 2 after the beep. <tone>
<record, transcribe, recognize (SpeechWorks)>

if silence, then

Please say the last name of person 2 AFTER you hear the beep. <tone>
<record, transcribe, recognize (SpeechWorks)>

Are there any other household members that need to be counted, yes or no? <tone>

if silence, then

Are there any other household members that need to be counted? Please say yes or no after the beep. <tone>

...(repeat for all members of household)

#####
PERSON 1
#####
[chime]
if roster > 1, then

For each of the persons on your list, we will now ask you a series of questions starting with yourself.

else

We will now ask you a series of questions about yourself.

#####NAME

:FN
Please spell your first name after the beep. <tone>
<record, transcribe, recognize (SpeechWorks)>

If silence, then

Please spell your first name AFTER you hear the beep. <tone>
<record, transcribe, recognize (SpeechWorks)>

:LN
Please spell your last name after the beep.<tone>
<record, transcribe, recognize (SpeechWorks)>

If silence, then

Please spell your last name AFTER you hear the beep.<tone>
<record, transcribe, recognize (SpeechWorks)>

:MI
Please say your middle initial. If there is no middle initial, say "none." Answer after the beep. <tone>
<record, transcribe, batch recognition later>

If silence,

Please tell us your middle initial. If there is no middle initial say "none". Answer AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

#####SEX

What is your sex, female or male? <tone>
<record, transcribe, batch recognition later>

if silence, then

What is your sex? Please answer either female or male AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

#####AGE & DATE_OF_BIRTH

[chime] We will now ask about your age and date of birth.

#####AGE
(April 1, 2000 and earlier)
What is your age on April 1, 2000? Please answer after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then

What is your age on April 1, 2000? Please answer AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

(April 2, 2000 and later)
What was your age on April 1, 2000? Please answer after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then

What was your age on April 1, 2000? Please answer AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

Please tell us the month, day, and year of your birth after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then

Please tell us the month, day, and year of your birth. Please answer AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

#####ORIGIN

[chime] We will now ask about your origin.

Are you of Spanish or Hispanic origin, yes or no? <tone>

if silence, then
Are you of Spanish or Hispanic origin? Please say yes or no after the beep. <tone>

<if yes, then>
Which of the following best describes your origin:
Mexican, Mexican American, Chicano, Puerto Rican, Cuban or other? <tone>

if silence, then
Which of the following best describes your origin:
Mexican, Mexican American, Chicano, Puerto Rican, Cuban or other? Please answer AFTER you hear the beep. <tone>

<if unrecognized, then go to OH>
<if other, then>
Okay, to what other Spanish or Hispanic group do you belong? <tone>
<record, transcribe, batch recognition later>

if silence, then
Please say what other Spanish or Hispanic group you consider yourself a member AFTER the beep. <tone>
<record, transcribe, batch recognition later>

<Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, other Hispanic>

:OH
Please spell that after the beep.<tone>
<record, transcribe, batch recognition later>

if silence, then
Please spell the name of the Spanish or Hispanic group AFTER the beep. <tone>
<record, transcribe, batch recognition later>

Do you belong to any other Spanish or Hispanic groups, yes or no? <tone>

if silence, then

Do you belong to any other Spanish or Hispanic groups? Please say yes or no after the beep. <tone>

<if yes, then>

Please tell us the name or names of these groups after the beep.<tone>
<record, transcribe, batch recognition later>
<Mexican, Mexican American, Chicano, Puerto Rican, Cuban, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, other Hispanic>

if silence, then

Please tell us the name or names of any other Spanish or Hispanic group you consider yourself a member AFTER the beep. <tone>
<record, transcribe, batch recognition later>

#####RACE

[chime] We will now ask about your race.

<If panel 2 or 5, then>

Do you belong to one or more of the races printed under question 8 on page 1 of the questionnaire, yes or no? <tone>

<If panel 8A or 8B, then>

Do you belong to one or more of the races printed under item 9 inside the brochure, yes or no? <tone>

if silence, then

goto RACELIST 1

<If yes, then>

Please say the name of the race or races you belong to with a short pause between each name after the beep.<tone>
<record, transcribe, batch recognition later>

if silence, then

goto RACELIST1

else

goto CONFIRM 1

<If no, then>

Please say the name of the other race or races you belong to with a short pause between each name after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then
goto RACELIST1
else
goto CONFIRM 1

:RACELIST1

Are you White? Please say yes or no after the beep. <tone>

<if yes, then>

Do you belong to any other races, yes or no? <tone>

if silence, then
Do you belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 1

Are you Black, African American, or Negro, yes or no? <tone>

if silence, then
Are you Black, African American, or Negro? Please say yes or no after the beep. <tone>

<if yes, then>

Do you belong to any other races, yes or no? <tone>

if silence, then
Do you belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 1

Are you an American Indian or Alaska Native, yes or no? <tone>

if silence, then

Are you an American Indian or Alaska Native? Please say yes or no after the beep. <tone>

<if yes, then>

Please say the name of your tribe after the beep. <tone>
<record, transcribe, batch recognition later>
<Cherokee, Blackfoot, Navajo, Chickasaw, Chippewa, Potawatomi, Sioux, Tohono O'Odham, Choctaw, Pima, Pueblo,Tlingit, Apache, Seminole, Iroquois, Alaskan Athabaskans, Lumbee, Cheyenne, Creek, Comanche, other tribe>

if silence, then

Please say the name of your tribe AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

Please spell the name of your tribe after the beep.<tone>
<record, transcribe, batch recognition later>

if silence, then

Please spell the name of your tribe AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

Do you belong to any other races, yes or no? <tone>

if silence, then

Do you belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>

skip to CONFIRMATION 1

Are you an Asian Indian, yes or no? <tone>

if silence, then

Are you an Asian Indian? Please say yes or no after the beep. <tone>

<if yes, then>

Do you belong to any other races, yes or no? <tone>

if silence, then

Do you belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>

skip to CONFIRMATION 1

Are you Chinese, yes or no? <tone>

if silence, then

Are you Chinese? Please say yes or no after the beep. <tone>

<if yes, then>

Do you belong to any other races, yes or no? <tone>

if silence, then

Do you belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>

skip to CONFIRMATION 1

Are you Filipino, yes or no? <tone>

if silence, then

Are you Filipino? Please say yes or no after the beep. <tone>

<if yes, then>

Do you belong to any other races, yes or no? <tone>

if silence, then

Do you belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>

skip to CONFIRMATION 1

Are you Japanese, yes or no? <tone>

if silence, then

Are you Japanese? Please say yes or no after the beep. <tone>

<if yes, then>

Do you belong to any other races, yes or no? <tone>

if silence, then

Do you belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>

skip to CONFIRMATION 1

Are you Korean, yes or no? <tone>

if silence, then

Are you Korean? Please say yes or no after the beep. <tone>

<if yes, then>

Do you belong to any other races, yes or no? <tone>

if silence, then

Do you belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>

skip to CONFIRMATION 1

Are you Vietnamese, yes or no? <tone>

if silence, then

Are you Vietnamese? Please say yes or no after the beep. <tone>

<if yes, then>

Do you belong to any other races, yes or no? <tone>

if silence, then

Do you belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>

skip to CONFIRMATION 1

Are you from some other Asian race, yes or no? <tone>

if silence, then
Are you from some other Asian race? Please say yes or no after the beep. <tone>

<if yes, then>
Please say the name of your race after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then
Please say the name of your race AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

<Cambodian, Hmong, Indonesian, Pakistani, Laotian, Thai, other Asian race>

Please spell the name or your race after the beep.<tone>
<record, transcribe, batch recognition later>

if silence, then
Please spell the name of your race AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

Do you belong to any other races, yes or no? <tone>

if silence, then
Do you belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 1

Are you Native Hawaiian, yes or no? <tone>

if silence, then

Are you Native Hawaiian? Please say yes or no after the beep. <tone>

<if yes, then>
Do you belong to any other races, yes or no? <tone>

if silence, then
Do you belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 1

Are you Guamanian or Chamorro, yes or no? <tone>

if silence, then
Are you Guamanian or Chamorro? Please say yes or no after the beep. <tone>

<if yes, then>
Do you belong to any other races, yes or no? <tone>

if silence, then
Do you belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 1

Are you Samoan, yes or no? <tone>

if silence, then
Are you Samoan? Please say yes or no after the beep. <tone>

<if yes, then>
Do you belong to any other races, yes or no? <tone>

if silence, then
Do you belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>

skip to CONFIRMATION 1

Are you from some other Pacific Islander race, yes or no? <tone>

if silence, then
Are you from some other Pacific Islander race? Please say yes or no after the beep. <tone>

<if yes, then>
Please say the name of your race after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then
Please say the name of your race AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

<Fijian, Palauan, Tahitian, Tongan, other Pacific Islander>

Please spell the name of your race after the beep.<tone>
<record, transcribe, batch recognition later>

if silence, then
Please spell the name of your race AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

Do you belong to some other race, yes or no? <tone>

if silence, then
Do you belong to some other race? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 1
<if yes, then>
Please say the name of your race after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then

Please say the name of your race AFTER you hear the beep. <tone>

<record, transcribe, batch recognition later>

Please spell the name of your race after the beep.<tone>
<record, transcribe, batch recognition later>

if silence, then
Please spell the name of your race AFTER you hear the beep. <tone>

<record, transcribe, batch recognition later>

####CONFIRMATION 1
[chime] We will now summarize the information you recorded about yourself.

{Silent responses should be spoken as "blank”}

Name: <first name 1 & last name 1>.
Sex: <sex>
Birthdate: if <month> = silence AND <day> = silence AND <year> = silence, then
say "blank”
else
<month> <day> <year>
Age: <age>
Origin: <origin> (if blank, say "Non-Hispanic”)
Race: <race>
Ownership: <owned free and clear, owned with a mortgage, rented for cash, occupied with no rent> {NOTE: these phrases are spoken, not synthesized}

Is all of this information correct, yes or no? <tone>

if silence, then
Is all of this information correct? Please say yes or no after the beep. <tone>

<if no, then>
Please tell us which item or items were incorrect and provide the correct information for each one after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then

For each item you wish to correct, please tell us the item and the new information AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

#####
PERSON 2..n
#####

[chime] We will now ask you some questions about <first name n> <last name n>.

#####NAME

Please spell the first name of this person after the beep. <tone>
<record, transcribe, recognize (SpeechWorks)>

if silence, then
Please spell the first name of this person AFTER you hear the beep. <tone>
<record, transcribe, recognize (SpeechWorks)>

Please spell the last name of this person after the beep. <tone>
<record, transcribe, recognize (SpeechWorks)>

if silence, then
Please spell the last name of this person AFTER you hear the beep. <tone>
<record, transcribe, recognize (SpeechWorks)>

Please say their middle initial. If there is no middle initial, say "none". Answer after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then
Please say their middle initial. If there is no middle initial, say "none". Answer AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

#####RELATIVE

Is <first name n> <last name n> related to you, yes or no? <tone>

if silence, then

Is <first name n> <last name n> related to you? Please say yes or no after the beep. <tone>

if yes, then

relation = true

if panel 2 or 5, then

Which one of the items listed under question 2 on page 2 of the form describes the relationship between this person and yourself? <tone>

if panel 8A or 8B, then

Which one of the items listed under item 5 inside the brochure describes the relationship between this person and yourself? <tone>

if silence, then

Which describes this person's relationship to you,
husband or wife,
natural born son or daughter,
adopted son or daughter,
stepson or stepdaughter,
brother or sister,
father or mother,
grandchild, parent-in-law,
son or daughter-in-law, or other relative? Please answer AFTER you hear the beep. <tone>

if "daughter", then

:D1 **Is this person your natural born daughter, yes or no? <tone>**

if silence, then

Is this person your natural born daughter? Please say yes or no after the beep. <tone>

if yes, then goto SEX

if no, then

:D2 **Is this person your adopted daughter, yes or no? <tone>**

if silence, then

Is this person your adopted daughter? Please say yes or no after the beep. <tone>

if yes, then goto SEX

if no, then

:D3 **Is this person your stepdaughter, yes or no? <tone>**

if silence, then

Is this person your stepdaughter? Please say yes or no after the beep. <tone>

if yes, then goto SEX

if "son", then

:S1 **Is this person your natural born son, yes or no? <tone>**

if silence, then

Is this person your natural born son? Please say yes or no after the beep. <tone>

if yes, then goto SEX

if no, then

:S2 **Is this person your adopted son, yes or no? <tone>**

if silence, then

Is this person your adopted son? Please say yes or no after the beep. <tone>

if yes, then goto SEX

if no, then

:S3 **Is this person your stepson, yes or no? <tone>**

if silence, then

Is this person your stepson? Please say yes or no after the beep. <tone>

If yes, then goto SEX

if unrecognized, then goto OREL

if other OR (if D1, D2, D3 are no OR silent) OR (if S1, S2, S3 are no OR silent), then

Please say what other relationship this person has with you after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then

Please say what other relationship this person has with you AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

:OREL

Please spell the name of the relationship this person has with you after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then

Please spell the name of the relationship this person has with you AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

if no, then

if panel 2 or 5, then

Which one of the items listed under question 2 on page 2 of the form describes the association between this person and yourself? <tone>

if panel 8A or 8B, then

Which one of the items listed under item 5 inside the brochure describes the association between this person and yourself? <tone>

if silence, then

Please say which of the following best describes the association between yourself and this person: a roomer, boarder, foster child, housemate, roommate, unmarried partner, other? Answer AFTER you hear the beep. <tone>

if unrecognized, then goto OASS

if other, then

Please say what other association this person has with you after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then

Please say what other association this person has with you after the beep. Answer AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

:OASS

Please spell the name of the association this person has with you after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then

Please spell the name of the association this person has with you. Please answer AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

#####SEX

What is this person's sex, female or male? <tone>
<record, transcribe, batch recognition later>

if silence, then

What is this person's sex? Please answer female or male after the beep. <tone>

<record, transcribe, batch recognition later>

#####AGE & DATE_OF_BIRTH

[chime] We will now ask about their age and date of birth.

(April 1, 2000 and earlier)
What will this person's age be on April 1, 2000? Please answer after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then

What will this person's age be on April 1, 2000? Please answer AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

(April 2, 2000 and later)
What was this person's age on April 1, 2000? Please answer after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then
What was this person's age on April 1, 2000? Please answer AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

Please tell us the month, day, and year this person was born after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then
Please tell us the month, day, and year this person was born. Please answer AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

#####ORIGIN

[chime] We will now ask about their origin

Is this person of Spanish or Hispanic origin, yes or no? <tone>

if silence, then
Is this person of Spanish or Hispanic origin? Please say yes or no after the beep. <tone>

if yes, then
Which of the following best describes their origin: Mexican, Mexican-American, Chicano, Puerto Rican, Cuban, or other? <tone>

if silent, then
Which of the following best describes their origin: Mexican, Mexican-American, Chicano, Puerto Rican, Cuban, or other? Please answer AFTER you hear the beep. <tone>

if unrecognized, then goto OH2
if other, then

Okay, to what other Spanish or Hispanic group do they belong? <tone>

<record, transcribe, batch recognition later>

if silent, then

What other Spanish or Hispanic group does this person belong? Please answer AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

<Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, other Hispanic>
:OH2

Please spell that after the beep. <tone>
<record, transcribe, batch recognition later>

if silent, then

Please spell the name of the Spanish or Hispanic group AFTER the beep. <tone>
<record, transcribe, batch recognition later>

Does this person belong to any other Spanish or Hispanic groups, yes or no? <tone>

if silent, then

Does this person belong to any other Spanish or Hispanice groups? Please say yes or no after the beep. <tone>

<if yes, then>

Please tell us the name or names of these groups after the beep. <tone>
<record, transcribe, batch recognition later>

if silent, then

Please tell us the name or names of any other Spanish or Hispanic group this person belongs to AFTER the beep. <tone>
<record, transcribe, batch recognition later>

<Mexican, Mexican American, Chicano, Puerto Rican, Cuban, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, other Hispanic>

#####RACE

[chime] We will now ask about their race.

<If panel 2 or 5, then>

Does this person belong to one or more of the races printed under question 8 on page 1 of the questionnaire, yes or no? <tone>

if silence goto RACE2

<If panel 8A or 8B, then>

Does this person belong to one or more of the races printed under item 9 inside the brochure, yes or no? <tone>

if silence goto RACE2

<If yes, then>

Please say the name of the race or races this person belongs to with a short pause between each name. <tone>
<record, transcribe, batch recognition later>

if silence, then
goto RACE2
else
goto CONFIRMATION 2

<If no, then>

Please say the name of the other race or races this person belongs to with a short pause between each name. <tone>
<record, transcribe, batch recognition later>

if silence, then
goto RACE2
else
goto CONFIRMATION 2

:RACE2

Is this person White? Please say yes or no after the beep. <tone>

<if yes, then>

Does this person belong to any other races, yes or no? <tone>

if silence, then

Does this person belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 2

Is this person Black, African American, or Negro, yes or no? <tone>

if silence, then
Is this person Black, African American, or Negro? Please say yes or no after the beep. <tone>

<if yes, then>
Does this person belong to any other races, yes or no? <tone>

if silence, then
Does this person belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 2

Is this person an American Indian or Alaska Native, yes or no? <tone>

if silence, then
Is this person American Indian or Alaskan Native? Please say yes or no after the beep. <tone>

<if yes, then>
Please tell us the name of this person's tribe after the beep. <tone>
<record, transcribe, batch recognition later>

<Cherokee, Blackfoot, Navajo, Chickasaw, Chippewa, Potawatomi, Sioux, Tohono O'Odham, Choctaw, Pima, Pueblo, Tlingit, Apache, Seminole, Iroquois, Alaskan Athabaskans, Lumbee, Cheyenne, Creek, Comanche, other tribe>

if silence, then
Please tell us the name of this person's tribe. Please answer AFTER you hear the beep. <tone>

<record, transcribe, batch recognition later>

Please spell the name of this person's tribe after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then
Please spell the name of this person's tribe AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

Does this person belong to any other races, yes or no? <tone>

if silence, then
Does this person belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 2

Is this person an Asian Indian, yes or no? <tone>

if silence, then
Is this person an Asian Indian? Please say yes or no after the beep. <tone>

<if yes, then>
Does this person belong to any other races, yes or no? <tone>

if silence, then
Does this person belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 2

Is this person Chinese, yes or no? <tone>

if silence, then
Is this person Chinese? Please say yes or no after the beep. <tone>

<if yes, then>

Does this person belong to any other races, yes or no? <tone>

if silence, then
Does this person belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 2

Is this person Filipino, yes or no? <tone>

if silence, then
Is this person Filipino? Please say yes or no after the beep. <tone>

<if yes, then>
Does this person belong to any other races, yes or no? <tone>

if silence, then
Does this person belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 2

Is this person Japanese, yes or no? <tone>

if silence, then
Is this person Japanese? Please say yes or no after the beep. <tone>

<if yes, then>
Does this person belong to any other races, yes or no? <tone>

if silence, then
Does this person belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 2

Is this person Korean, yes or no? <tone>

if silence, then

Is this person Korean? Please say yes or no after the beep. <tone>

<if yes, then>

Does this person belong to any other races, yes or no? <tone>

if silence, then

Does this person belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>

skip to CONFIRMATION 2

Is this person Vietnamese, yes or no? <tone>

if silence, then

Is this Vietnamese? Please say yes or no after the beep. <tone>

<if yes, then>

Does this person belong to any other races, yes or no? <tone>

if silence, then

Does this person belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>

skip to CONFIRMATION 2

Is this person from some other Asian race, yes or no? <tone>

if silence, then

Is this person from some other Asian race? Please say yes or no after the beep. <tone>

<if yes, then>

Please say the name of this person's race after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then

Please say the name of this person's race AFTER you hear the beep. <tone>

<record, transcribe, batch recognition later>

<Cambodian, Hmong, Indonesian, Pakistani, Laotian, Thai, other Asian>

Please spell the name of this person's race after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then
Please spell the name of this person's race AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

Does this person belong to any other races, yes or no? <tone>

if silence, then
Does this person belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 2

Is this person Native Hawaiian, yes or no? <tone>

if silence, then
Is this person Native Hawaiian? Please say yes or no after the beep. <tone>

<if yes, then>
Does this person belong to any other races, yes or no? <tone>

if silence, then
Does this person belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>
skip to CONFIRMATION 2

Is this person Guamanian or Chamorro, yes or no? <tone>

if silence, then

Is this person Guamanian or Chamorro? Please say yes or no after the beep. <tone>

<if yes, then>

Does this person belong to any other races, yes or no? <tone>

if silence, then

Does this person belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>

skip to CONFIRMATION 2

Is this person Samoan, yes or no? <tone>

if silence, then

Is this person Samoan? Please say yes or no after the beep. <tone>

<if yes, then>

Does this person belong to any other races, yes or no? <tone>

if silence, then

Does this person belong to any other races? Please say yes or no after the beep. <tone>

<if no, then>

skip to CONFIRMATION 2

Is this person from some other Pacific Islander race, yes or no? <tone>

if silence, then

Is this person from some other Pacific Islander race? Please say yes or no after the beep. <tone>

<if yes, then>

Please say the name of this person's race after the beep. <tone>
<record, transcribe, batch recognition later>

if silence, then

Please say the name of this person's race AFTER you hear the beep. <tone>

<record, transcribe, batch recognition later>

<Fijian, Palauan, Tahitian, Tongan, other Pacific Islander>

Please spell the name of this person's race after the tone. <tone>
<record, transcribe, batch recognition later>

if silence, then

Please spell the name of this person's race AFTER you hear the beep. <tone>

<record, transcribe, batch recognition later>

Is this person from some other race, yes or no? <tone>

if silence, then

Is this person from some other race? Please say yes or no after the beep. <tone>

<if yes, then>

Please say the name of this person's race. <tone>
<record, transcribe, batch recognition later>

if silence, then

Please say the name of this person's race AFTER you hear the beep. <tone>

<record, transcribe, batch recognition later>

Please spell the name of this person's race after the tone. <tone>
<record, transcribe, batch recognition later>

if silence, then

Please spell the name of this person's race AFTER you hear the beep. <tone>

<record, transcribe, batch recognition later>

####CONFIRMATION 2
[chime] We will summarize the information you recorded about this person.

{Silent responses should be spoken as "blank"}

Name: <first name n> <last name n>.
Sex: <sex>

Birthdate: if <month> = silence AND <day> = silence AND <year> = silence, then
say "blank"
else
<month> <day> <year>
Age: <age>
if relation, then
Relationship: <relationship>
(if initial response was ambiguous, i.e. son or daughter, then the applicable phrase should be announced here. Choose from
adopted son, natural born son, stepson, adopted daughter, natural born daughter, and stepdaughter
else
Association: <association>
Origin: <origin> (if blank, say "Non-Hispanic")
Race: <race>

Is all of this information correct, yes or no? <tone>

if silence, then
Is all of this information correct? Please answer yes or no after the beep. <tone>

<if no, then>
Please tell us which item or items were incorrect and provide the correct information after each one. <tone>
<record, transcribe, batch recognition later>

if silence, then
For each item you wish to correct, please tell us the item and the new information AFTER you hear the beep. <tone>
<record, transcribe, batch recognition later>

If more people, then go to PERSON 2..n

##CHECK FOR COMPLETENESS
complete = true
For person 1 to roster, do
if roster = 1, then
2. HOME_OWNER –one answer category;

3. **Name--first and last name fields are treated as one item and together must have a minimum of three alpha characters; middle initial is not considered;**
4. **Sex–one answer category;**
5. **(Age or Date of Birth) OR (Age or year of birth) OR (Age or month and day of birth);**
6. **Hispanic Origin–at least one answer category;**
7. **Race--at least one answer category.**

if 5 out of 6 of the conditions above are false, then
if 5 out of 6 fields above are silent, then
complete = false

if roster > 1, then for each person 2..n,

1. **Relationship–one answer category;**
2. **Name--first and last name fields are treated as one item and together must have a minimum of three alpha characters; middle initial is not considered;**
3. **Sex–one answer category;**
4. **(Age or Date of Birth) OR (Age or year of birth) OR (Age or month and day of birth);**
5. **Hispanic Origin–at least one answer category;**
6. **Race--at least one answer category.**

if 5 out of 6 of the conditions above are false, then
if 5 out of 6 fields above are silent, then
complete = false

if not complete, then
[chime]
if panel 5 or 8A, then
We did not receive enough information from you to activate your calling card.
else
Your Census form is not complete. If you need help....

Please call us at 1-877-8-CENSUS for assistance. A Census worker may contact you later to complete the rest of your information.
goto GOODBYE

[chime] We will now take your calling card information.
If panel 5, then
Please remove the calling card from the insert and turn it over to see the calling card number.

If panel 8A, then

Please remove the calling card from the brochure and turn it over to see the calling card number.

<If panel 8A, then>

<lookup calling card number>

Our records show that the calling card we sent you has this ID: xxxxxxxxxxx.

Is this correct, yes or no? <tone>

if silence, then

Is this correct? Please say yes or no after the beep. <tone>

If yes, then

This card is valid for one year and you may begin using it immediately.

<skip to SURVEY QUESTIONS>

If no, then

If buttons, then

Please enter all ten digits of the calling card number with the pushbutton keys on your telephone now.<tone2>

else

Please say all ten digits of the calling card number now. <tone>

<lookup CC #>

go to NO MATCH

<If panel 5, then>

<If buttons, then>

Please enter all ten digits of your calling card number with the pushbutton keys on your telephone now. <tone>

<else>

Please say the all ten digits of your calling card number now. <tone>

<lookup CC #>

##NO MATCH

<If no match, then>

If buttons, then

The number you entered is not in our records, please enter all ten digits of your calling card number again. <tone>

else

The number you entered is not in our records, please say all ten digits of your calling card number again. <tone>

<else>

This card is valid for one year and you may begin using it immediately.
<skip to SURVEY QUESTIONS>

<lookup CC #>

<If no match, then>

See bailout specification

<else>

This card is valid for one year and you may begin using it immediately.

#####SURVEY QUESTIONS
[chime] We will now ask you some questions about this system.

On a scale of 1 to 5, where 5 means Very Satisfied, 1 means Very Dissatisfied, and 3 means neither Satisfied nor Dissatisfied, how Satisfied are you overall with the computerized questionnaire? <tone>

if silence, then

On a scale of 1 to 5, where 5 means Very Satisfied, 1 means Very Dissatisfied, and 3 means neither Satisfied nor Dissatisfied, how Satisfied are you overall with the computerized questionnaire? Please answer AFTER you hear the beep. <tone>

<score = 1..5>

<if score < 3, then>

Please tell us what you disliked about the computerized questionnaire after the beep. <tone>
<record, transcribe>

if silence, then

Please tell us what you disliked about the questionnaire AFTER you hear the beep. <tone>
<record, transcribe>

<if score > 3, then>
Please tell us what you liked about the computerized questionnaire after the beep. <tone>
<record, transcribe>

if silence, then
Please tell us what you liked about the questionnaire AFTER you hear the beep. <tone>
<record, transcribe>

Were you able to fully understand the computer, yes or no? <tone>

if silence, then
Were you fully able to understand the computer? Please answer yes or no after the beep. <tone>

<if no, then>
Please tell us what you did not understand after the beep. <tone>
<record, transcribe>

if silence, then
Please tell us what you did not understand AFTER you hear the beep. <tone>
<record, transcribe>

Was the computer able to fully understand you, yes or no? <tone>

if silence, then
Was the computer fully able to understand you? Please say yes or no after the beep. <tone>

<if no, then>
Please tell us what the computer did not understand after the beep. <tone>
<record, transcribe>

if silence, then

Please tell us what the computer did not understand AFTER you hear the beep. <tone>
<record, transcribe>

Was there anything about the questionnaire that was confusing or frustrating, yes or no? <tone>

if silence, then

Was there anything about the questionnaire that was confusing or frustrating? Please answer yes or no after the beep. <tone>

<if yes, then>

Please tell us what was confusing or frustrating after the beep. <tone>
<record, transcribe>

if silence, then

Please tell us what was confusing or frustrating AFTER you hear the beep. <tone>
<record, transcribe>

Did you have too much time, too little time, or just the right amount of time to answer the questions? <tone>
<record, transcribe>

if silence, then

Did you have too much time, too little time, or just the right amount of time to answer the questions? Please answer AFTER you hear the beep. <tone>
<record, transcribe>

Please tell us your suggestions about improving the computerized questionnaire after the beep. <tone>
<record, transcribe>

if silence, then

Please tell us your suggestions about improving the questionnaire AFTER you hear the beep. <tone>
<record, transcribe>

#####GOODBYE
[chime]

Thank you for your help with the 2000 Census. You do not need to mail in your paper questionnaire. Goodbye.

Appendix C

Design of the Automated Spoken Questionnaire System

An Automated Spoken Questionnaire (ASQ) system has several components:

- **Speech recognition software;**
- **Automated digital signal processor boards;**
- **Database management software;**
- **Survey software; and**
- **Telephony software.**

Speech recognition software allows the system to accept any respondent's spoken replies. Each time the respondent speaks, the software refers to a special file called a "context file" that lists the words that the respondent might use to reply. For example, the context file for a yes-or-no question might contain "yes," "no," "yeah," "no way," "certainly," and so on. The speech recognition software identifies the best match to the respondent's utterance. It assigns a level of confidence, expressed as a percentage, to the match.

In this project, the speech recognition software was from Lernout and Hauspie.

An Automated Digital Signal Processor board serves as the interface between the computer and the telephone. This hardware transforms the respondent's voice into a signal that the computer can treat as incoming data. It also allows the computer to play digitized voice files to the respondent. In this project, hardware for the ASQ application included a single processor 200 MHz NT server equipped with a Dialogic Antares Digital Speech Processing board.

When this system recorded the respondents' spoken replies, it stored them as .wav files (PCM 8,000 Hz, 8 bit mono, 8K/second). The Antares board was programmed to end-point the speech files (i.e., determine that the respondent had finished speaking and stop recording) after 2 seconds of silence. Barge-in (i.e., the capability of the respondent to begin speaking at any time, even while the system was still playing a voice file) was not enabled anywhere in the questionnaire.

Prior to the start of this study, the Census Bureau provided Westat with a mathematical model to determine the number of telephone lines that would be sufficient to handle the maximum call volume expected to the ASQ system during this study. Although that mathematical model suggested that seven lines would suffice, Westat provided eight lines in case the model was too conservative.

Database Management software allows the computer to store the responses and transfer them via the Internet to a central database. In this study, the ASQ system stored data using Microsoft Excel.

Survey software contains the programming that directs the flow of the questionnaire. This software selects the voice files that the computer must play and obtains the responses to the questions. Westat staff wrote the software for this census questionnaire application in Visual C++.

Telephony software enables ASQ respondents to speak to an operator when needed. For example, some respondents might have unusual voices which the speech recognition software cannot *process*. In that situation, telephony software may connect these respondents with an operator who would administer the survey. The telephony software for this project was the CallCenter@nywhere product of Telephony@Work.

Appendix D

CENSUS 2000

TELEPHONE INTERVIEWING SPECIFICATIONS FOR SHORT FORM

SPECS: All numeric fields must be stored as right justified. All character fields must be stored as left justified.

Programming Note: Check formtype from TQA and follow appropriate path. If formtype is missing, then default to D10 (No ID path).

Refer to the last page for a definition of completed interviews.

D-1 = Interview with a 22 character census ID number.
D-10 = Interview without a 22 character census ID Number, assign processing ID as noted below and start at >address_a<

Assignment of processing id:
Character 1-2 66 = TQA generated BCF interviews
Character 3-5 000 = English
001 = English (in sequence for 000 is filled)
002 = Spanish
003 = Chinese
004 = Korean
005 = Tagalog
006 = Vietnamese
007 = English (PR)
008 = Spanish (PR)

Character 6-12 sequence number 0000001 - 9999999

Character 13-14 MAD97 check digits

SPECS: If census identification number was forwarded from TQA, skip to >POP_count<using formtype to indicate which form to complete. If NO census identification number was provided, skip to >ID<. The census identification number must be part of the output information for the Bureau of the Census with the short form data from the interviews.

>ID<

If you have your census form available, please refer to the census identification number located on the back page underneath the bar code. What is the ID number on your questionnaire?

_ _ _ _ _-_ _ _ _ _ _ _-_ _-_ _ _-_ _ _-_ _ (allow 22 characters)--Skip to >POP_count<

Programming Note: If ID entered, perform check digit algorithm. If check fails increment ckdig2 counter by 1. If ckdig2=3, then blank ID and go to >address_a<

(N) Not available,

If no ID number, set formtype=D10 and go to >address_a<;
If valid ID, and Phone Num. available, go to >ANIchk<,
If valid ID, and No Phone Num., go to >GetPhone<

OUTPUT SPECS: Store in Record 1, at location 2.

Programming note: If uheflag=1, use second fill. Else, use first fill.

>address_a<

We need to be sure that everyone is counted correctly in the census. I'd like to take some information about your household, starting with your home address.
What is the mailing address where you (lived on Saturday, April 1, 2000/ live or stay MOST OF THE TIME)?

Mailing Address: _____ >housestreet1< (allow 63 characters—If address is a P.O. Box address, Rural Route/Box address or No address, store address collected by agent starting at character 9 so that 8-letter string may be stored in characters 1-8. See specs below. For house number and street/road name style address, start address at character 1.)

SPECS: P.O. Box address, Rural Route/Box address or No address are stored in >housestreet1<.

AND CHECK ONE BOX IF APPROPRIATE:

[ ] P.O. Box address -- Skip to >aptno1<

SPECS: For output, fill >housestreet1< to position 62. At position 63, fill with a "P." Set nohouse = 1.

[ ] Rural Route/Box address – Skip to >aptno1<

SPECS: For output, fill >housestreet1< to position 62. At position 63, fill with an "R.". Set nohouse = 1

[ ] No address on April 1, or address is a location description such as a park, or street names – Skip to >zip1<

SPECS: For output, fill >housestreet1< to position 62. At position 63, fill with an "O." Also, if "No address" box is marked, store "1" in variable bcmailno at location 81. Set nohouse = 1.

OUTPUT SPECS: Store in Record 1 for D-10 at the following locations:
bcmailno: location 81
housestreet1: location 82

Programming note: Do NOT advance to the next screen when write-in field is blank unless the "No address" box is marked.
_____

>aptno1<

Do you have an apartment number?

Yes -- _____ Apartment number (allow 16 characters)
No

OUTPUT SPECS: Store apartment number in Record 1 for D-10 at location 145.
_____

>zip1<

What is the ZIP code?

_____ (allow 5 characters)

OUTPUT SPECS: Store in Record 1 for D-10 at location 195.
_____

>city1<

What is the name of your city or town?

_____ (allow 16 characters)

OUTPUT SPECS: Store in Record 1 for D-10 at location 161.
_____

>state1<

What state? SELECT THE STATE USING THE ARROW KEYS IF NECESSARY AND PRESS THE ENTER KEY.

_____ (allow 2 characters)

(1) AL--Alabama
(2) AK--Alaska
(3) AZ--Arizona
(4) AR--Arkansas
(5) CA--California
(6) CO--Colorado
(7) CT--Connecticut
(8) DE--Delaware
(9) DC--District of Columbia
(10) FL--Florida
(11) GA--Georgia
(12) HI--Hawaii
(13) ID--Idaho
(14) IL--Illinois
(15) IN--Indiana
(16) IA--Iowa
(17) KS--Kansas
(18) KY--Kentucky
(19) LA--Louisiana
(20) ME--Maine
(21) MD--Maryland
(22) MA--Massachusetts
(23) MI--Michigan
(24) MN--Minnesota
(25) MS--Mississippi
(26) MO--Missouri
(27) MT--Montana
(28) NE--Nebraska
(29) NV--Nevada
(30) NH--New Hampshire
(31) NJ--New Jersey
(32) NM--New Mexico
(33) NY--New York
(34) NC--North Carolina
(35) ND--North Dakota
(36) OH--Ohio
(37) OK--Oklahoma
(38) OR--Oregon
(39) PA--Pennsylvania
(40) RI--Rhode Island
(41) SC--South Carolina
(42) SD--South Dakota
(43) TN--Tennessee
(44) TX--Texas
(45) UT--Utah
(46) VT--Vermont
(47) VA--Virginia
(48) WA--Washington
(49) WV--West Virginia
(50) WI--Wisconsin
(51) WY--Wyoming

SPECS: Do NOT output codes. Output 2-letter abbreviation associated with codes.

OUTPUT SPECS: Store in Record 1 for D-10 at location 193.

>county1<

What county is that city or town in?

_____ (allow 16 characters)

D Don't know
R Refused

SPECS: If nohouse=1 then go to >bchsnnno<, else go to >bcallres<

OUTPUT SPECS: Store in Record 1 for D-10 at location 177.

Programming note: From the Mailing Address screen, do not advance to the next screen without the city and state.
_____

SPECS: Ask if nohouse=1. Else, skip to >bcallres<

>bchsnnno<

Do you have a street address with a house number?

(1) Yes -- Skip to >housestreet2<
(2) No – Skip to >bcallres<

OUTPUT SPECS: If 2, store "1" in Record 1 for D-10 at location 200, else leave this location blank.

Programming note: Do NOT advance to the next screen without a Yes or No answer.
_____

>housestreet2<

House number and street/road name _____ (allow 34 characters)

OUTPUT SPECS: Store in Record 1 for D-10 at location 201.
_____

>aptno2<

Do you have an apartment number?

Yes -- _____ Apartment number (allow 16 characters)
No

OUTPUT SPECS: Store apartment number in Record 1 for D-10 at location 235.

>zip2<

What is the ZIP code?

_____ (allow 5 characters)

OUTPUT SPECS: Store in Record 1 for D-10 at location 285.

>city2<

What is the name of your city or town?

_____ (allow 16 characters)

OUTPUT SPECS: Store in Record 1 for D-10 at location 251.

>state2<

What state? SELECT THE STATE USING THE ARROW KEYS IF NECESSARY AND PRESS THE ENTER KEY.

_____ (allow 2 characters)

(1) AL--Alabama
(2) AK--Alaska
(3) AZ--Arizona
(4) AR--Arkansas
(5) CA--California
(6) CO--Colorado
(7) CT--Connecticut
(8) DE--Delaware
(9) DC--District of Columbia
(10) FL--Florida
(11) GA--Georgia
(12) HI--Hawaii
(13) ID--Idaho
(14) IL--Illinois
(15) IN--Indiana
(16) IA--Iowa
(17) KS--Kansas
(18) KY--Kentucky
(19) LA--Louisiana
(20) ME--Maine
(21) MD--Maryland
(22) MA--Massachusetts
(23) MI--Michigan
(24) MN--Minnesota
(25) MS--Mississippi
(26) MO--Missouri
(27) MT--Montana
(28) NE--Nebraska
(29) NV--Nevada
(30) NH--New Hampshire
(31) NJ--New Jersey
(32) NM--New Mexico
(33) NY--New York
(34) NC--North Carolina
(35) ND--North Dakota
(36) OH--Ohio
(37) OK--Oklahoma
(38) OR--Oregon
(39) PA--Pennsylvania
(40) RI--Rhode Island
(41) SC--South Carolina
(42) SD--South Dakota
(43) TN--Tennessee
(44) TX--Texas
(45) UT--Utah
(46) VT--Vermont
(47) VA--Virginia
(48) WA--Washington
(49) WV--West Virginia
(50) WI--Wisconsin
(51) WY--Wyoming

SPECS: Do NOT output codes. Output 2-letter abbreviation associated with codes.

OUTPUT SPECS: Store in Record 1 for D-10 at location 283.

>county2<

What county is that city or town in?

_____ (allow 16 characters)

D Don't know
R Refused

SPECS: Go to >bcallres<

OUTPUT SPECS: Store in Record 1 for D-10 at location 267.

Programming note: From the Second Address screen, do NOT advance to the next screen without the city and state.

_____

>bcallres<

Programming Note: Ask if formtype=D10; if uheflag=1 use second fill, else use first fill.

(Are you calling to complete an interview for <u>all</u> the people who were living or staying at this address on April 1, 2000?/ Are you calling to complete an interview for <u>all</u> the people who were living or staying at the place where you live MOST OF THE TIME?)

(1) Yes
(2) No

SPEC: If Phone Num. available, go to >ANIchk<
If No Phone Num., go to >GetPhone<

OUTPUT SPECS: Store in Record 1 for D-10 at location 290

SPECS: Ask if Phone Num. available, else go to >GetPhone<

>ANIchk<

We used our caller ID system to capture the phone number you are calling from. Is (FILL with verified/corrected ANI) your correct home phone number?

(1) Yes -- Store as >phonarea<, >phonepre<, >phonesfx<
(2) No -- read:
What is your home phone number starting with your area code?

_____ **
(area (prefix) (suffix) code)

****Note: Phone number is captured as one field, but output as three fields.**

SPECS: If formtype=D-1, go to >POP_count<
If formtype=D-10 and:
>bcallres< =1 then go to >POP_count<
>bcallres< =2 then go to >tenure<

OUTPUT SPECS: Store the area code as >phonarea<; store the prefix as >phonepre<; store the suffix as >phonesfx<.

Store in Record 1 as follows:

	D-10	D-1
>phonarea<:	location 291	location 272
>phonepre<:	location 294	location 275
>phonesfx<:	location 297	location 278

>GetPhone<

SPEC: Ask if No Phone Num.

In case we need to contact you later, please give me your home phone number starting with your area code.

_____ **
(area (prefix) (suffix) code)

_ Don't know/Refused

****Note: Phone number is captured as one field, but output as three fields.**

SPECS: If formtype=D-1, go to >POP_count<
If formtype=D-10 and:

>bcallres< =1 then go to >POP_count<
>bcallres< =2 then go to >tenure<

OUTPUT SPECS: Store the area code as >phonarea<; store the prefix as >phonepre<; store the suffix as >phonesfx<.

Store in Record 1 as follows:

	D-10	D-1
>phonarea<:	**location 291**	**location 272**
>phonepre<:	**location 294**	**location 275**
>phonesfx<:	**location 297**	**location 278**

SPECS: Include the residence rules job aid from knowledge data base as HELP.

>POP_count<

Programming note: If uheflag=1 use second fill, else use first fill.

(How many people were living or staying in this house, apartment, or mobile home on April 1, 2000/ How many people were living or staying in the house, apartment, or mobile home where you live or stay MOST OF THE TIME)?

___ (allow 2 characters)

(H) HELP

OUTPUT SPECS: Store in Record 1 for D-1 or D-10, at location 83.

>tenure<

Programming note: If uheflag=1 use second fill, else use first fill

(Is this house, apartment, or mobile home—/ Is the house, apartment, or mobile home where you live or stay MOST OF THE TIME---)

(1) Owned by you or someone in this household with a mortgage or loan?
(2) Owned by you or someone in this household free and clear without a mortgage or loan?
(3) Rented for cash rent?
(4) Occupied without payment of cash rent?

D Don't know
R Refused

SPECS: If "D" or "R", store 0. If >bcallres< = 2, skip to >partial_roster<. Else, go to >roster<.

OUTPUT SPECS: Store for D-1 in Record 1, at location 282. Store for D-10, Record 1, at location 301.

>partial_roster<
This screen does not exactly reflect the OSS format.

Programming note: Ask if bcallres=2, else go to >roster<; If uheflag=1 use second fill, else use first fill.

(What are the names of the persons who were living or staying in this house, apartment, or mobile home on April 1, 2000/ What are the names of the persons who were living or staying in the house, apartment, or mobile home where you live or stay MOST OF THE TIME)? Start with yourself or a person living with you who was not counted.

ENTER NAMES

FIRST NAME	MIDDLE INITIAL	LAST NAME	CODE

(ALLOW ENOUGH SPACE FOR PEOPLE)
Allow 15 characters for last name.
Allow 13 characters for first name.
Allow 1 characters for middle initial.
Allow 1 character for code.

Add boxes for indicating "respondent" and "proxy."

OUTPUT SPECS: Store a "2" in location 1 for record type 2. Store in Record 2, for D-1 or D-10 at locations:
Last Name: location 47
First Name: location 63
Middle Initial: location 76

If code=A then PSTATUS=1, if code=D then PSTATUS=3; for D-1 store PSTATUS in record type 2, location 44.

Store a "1" in PNUM at location 42 for person on line 1 of the roster; store a "2" in PNUM at location 42 for person on line 2 of the roster; etc.

SPEC: If bcallres=2, then go to >resp@1<

>roster<
This screen does not exactly reflect the OSS format.

Programming note: if uheflag=1 use second fill, else use first fill.

What are the names of all persons who were living or staying (in this house, apartment, or mobile home on April 1, 2000/ at the house, apartment, or mobile home where you live or stay MOST OF THE TIME)? Start with the name of one of the people living here who owns, is buying, or rents this house, apartment, or mobile home.

ENTER NAMES

FIRST NAME	MIDDLE INITIAL	LAST NAME	CODE

(ALLOW ENOUGH SPACE FOR PEOPLE)
Allow 15 characters for last name.
Allow 13 characters for first name.
Allow 1 characters for middle initial.
Allow 1 character for code.

Add boxes for indicating "respondent" and "proxy."

OUTPUT SPECS: Store a "2" in location 1 for record type 2. Store in Record 2, for D-1 or D-10 at locations:
Last Name: location 47
First Name: location 63
Middle Initial: location 76

If code=A then PSTATUS=1, if code=D then PSTATUS=3; for D-1 store PSTATUS in record type 2, location 44.

Store a "1" in PNUM at location 42 for person on line 1 of the roster; store a "2" in PNUM at location 42 for person on line 2 of the roster; etc.

SPEC: If person added from >coverage1< then go to >coverage 2<, else proceed to >coverage1<.

>coverage1<

Programming note: If uheflag=1 use second fill, else use first fill.

Did anyone else such as housemates, roommates, live-in employees, boarders, foster children or anyone temporarily away on business or vacation live (at this address on April 1, 2000/ at the place where you live MOST OF THE TIME)?

(1) Yes--Ask: What is/are their name(s)?
(2) No

SPECS: If "1" go to >roster<. Add the name(s) to the list and enter an "A" in the "Code" column for each name given. If "2" go to >coverage2<. Increase the number in >POP_count< if names are added.

>coverage2<

Did you include any people who were living away at college, in the Armed Forces and living somewhere else, in a correctional facility, in a mental hospital, in a nursing home, hospice or ward for the chronically ill, or staying at another residence most of the week while working?

(1) Yes--Ask: What is/are their name(s)?
(2) No

SPECS: If "1" go to >roster<. Enter a "D" in the "Code" column beside each name given.
Decrease the number in >POP_count< if names are deleted.

>adc_names<

I'm going to read you the list of people to verify that all names are listed correctly. (READ NAMES AND VERIFY SPELLING)

MAKE SURE [fill with respondent's name] IS LISTED ON ROSTER
EXCEPT FOR A PROXY

(P) All correct
(A) Add person not listed
(D) Delete person listed
(C) Spelling Change
(U) Undelete person listed

ENTER LINE NUMBER OF PERSON: ___ (To advance to >Spelling< and take the

appropriate action above)

SPECS: If "A", increase >POP_count<. If "D", decrease >POP_count<.

INTERVIEWER SCREEN
>resp@1<
(Ask if necessary.)
Enter line number of person you are talking to ____

SPECS: We need to allow for proxy interviews (code 99). A "proxy interviewer" is someone providing the interview that is not on the >roster<.

>Spelling<

MAKE THE CORRECTIONS NEEDED:

First _____

MI _____

Last _____

SPECS: Ask the >relation< question of EVERYONE listed on the Roster except the person on line one. For person one, store 0 in >relation<. Ask the >otherrel< question only if the answer is "10". Then continue with >sex1< starting with the person on line one.

>relation<

How (are you /is ...) related to (fill with the name on line 1)*?

(1) Husband/wife
(2) Natural-born son/daughter
(3) Adopted son/daughter
(4) Stepson/stepdaughter
(5) Brother/sister
(6) Father/mother
(7) Grandchild
(8) Parent-in-law
(9) Son-in-law/daughter-in-law
(10) Other relative--Skip to >otherrel<
(11) Roomer, boarder
(12) Housemate, roommate
(13) Unmarried partner
(14) Foster child
(15) Other nonrelative

D Don't know
R Refused

Skip to >sex1< except for category (10).

*** If respondent is person on line 1, use "you" instead of name of person on line 1.**
SPECS: If "D" or "R", store 0.

OUTPUT SPECS: Store in Record 2 for D-1 or D-10, at location 77.

>otherrel<

SPECS: Ask only if needed or fill with information the respondent provided when asked >relation<.

How (are you/is...) related?

_____ (allow for 12 characters)

D Don't know
R Refused

SPECS: If "D" or "R", store a "Blank".

OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at location 79.

>sex1<

SPECS: FOR THE RESPONDENT SCREEN ONLY, ADD: ASK ONLY IF NECESSARY.
SPECS: Ask question of everyone listed on the Roster before continuing to >dob<.

(Are you/Is...) male or female?

(1) Male
(2) Female

D Don't know
R Refused

SPECS: If "D" or "R", store 0.

OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at location 91.

SPECS: Ask questions of EACH person on the roster before continuing to >hisp_origin1<.

>dob<

What is (your/...'s) date of birth?

Month	Day	Year of birth
_____	____	__/__/__/__ (allow 1884-2000)
(01-12)	(01-31)	
dob@mth	dob@dy	dob@yr

D Don't know--Skip to >age<
R Refused--Skip to >age<

SPECS: If any part (month, day, or year of birth) is Don't know or Refused, skip to >age<. For year of birth, output full year such as "1985" and not "985." If "D" or "R" in any field, store a "Blank". Output >dob@mth< as >DOBMONTH_4<; >dob@dy< as >DOBDAYXX_4<; and >dob@yr< as >DOBYEARX_4<.

OUTPUT SPECS: Store in Record 2, for D-1 or D-10 and the following locations:
Month: location 95
Day: location 97
Year of Birth: location 99

SPECS: If computed age is less than 1 year, substitute the "computed age" with appropriate "months/weeks". This screen does not exactly reflect the OSS format.

>ver_age<

So (were you/was...) (computed age) years old on April 1, 2000?

(1) Yes--Skip to >hisp_origin1<
(2) No

D Don't know--Skip to >hisp_origin1<
R Refused--Skip to >hisp_origin1<

SPECS: If 1, store computed age in >age<. If the computed age is less than 1 year, store "0". If "D" or "R", store a "Blank".

OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at location 92.

>age<

What was (your/...'s) age on April 1, 2000? IF CALLER DOES NOT KNOW THE EXACT AGE – Please estimate (your/…'s) age on April 1, 2000?

___ (allow 0-116)

D Don't know
R Refused

SPECS: Store age in >age<. If the computed age is less than 1 year, store "0". If "D" or "R", store a "Blank".

OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at location 92.

SPECS: Ask questions of EACH person on roster before continuing to >race<.

>hisp_origin1<

(Are you/Is ...) Spanish, Hispanic, or Latino? READ IF NECESSARY: FOR EXAMPLE, MEXICAN, MEXICAN AMERICAN, CHICANO, PUERTO RICAN, CUBAN, OR ANOTHER SPANISH, HISPANIC, OR LATINO GROUP.

(1) Yes – continue to >hisp_origin2<
(2) No--Skip to >race<

D Don't know--Skip to >race<
R Refused--Skip to >race<

SPECS: If 2, store 1 in >HISPCB01_5<. If "D" or "R", store 0.

OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at location 103.

>hisp_origin2<

SPECS: Accept only ONE response.

Which one of the following Spanish, Hispanic, or Latino groups (do you/does …) identify with? Mexican, Mexican American, Chicano, Puerto Rican, Cuban, or another Spanish, Hispanic or Latino group.

(1) Mexican, Mexican American, Chicano--Skip to >race<
(2) Puerto Rican--Skip to >race<
(3) Cuban--Skip to >race<
(4) Other Spanish/Hispanic/Latino--continue to >othr_sp1<

D Don't know--Skip to >race<
R Refused--Skip to >race<

SPECS: Store answers of 1, 2, 3, 4 as follows:

Current answer	Stored value	Variable
1	1	>HISPCB02_5<
2	1	>HISPCB03_5<
3	1	>HISPCB04_5<
4	1	>HISPCB05_5<

OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at the following locations:
HISPCB02_5: location 104
HISPCB03_5: location 105
HISPCB04_5: location 106
HISPCB05_5: location 107

>othr_sp1<

What is this group?

(1) Argentinean--Skip to >race<
(2) Colombian--Skip to >race<
(3) Dominican--Skip to >race<
(4) Nicaraguan--Skip to >race<
(5) Salvadoran--Skip to >race<
(6) Spaniard--Skip to >race<
(7) Other--Skip to >othr_sp<

D Don't know--Skip to >race<
R Refused--Skip to >race<

SPECS: Store words corresponding to categories 1 through 6 in >HISPANWI_5<.

OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at location 108.
_____
_____
>othr_sp<

What is the name of the other Hispanic group?

_____
(allow for 19 characters)

D Don't know
R Refused

SPECS: Store >othr_sp< in >HISPANWI_5<. If "D" or "R", store "Blank."

OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at location 108.
_____
_____

SPECS: Respondent may choose one or more categories.
SPECS: Ask questions of each person listed on the Roster in the following order:
>race<, >othr_race<, >amer_ind<, >asian_group<, >othr_asian<, >pacific_group<, and >othr_pacific< (when appropriate) before continuing to the next person.

>race<

I'm going to read a list of race categories. Please choose one or more categories that best indicate (your/...'s) race. (Are you/Is...) White? Black, African American or Negro? American Indian or Alaska Native? Asian? Native Hawaiian? Other Pacific Islander? or Some other race?

(1) White
(2) Black, African American, or Negro
(3) American Indian or Alaska Native--Skip to >amer_ind<
(4) Asian--Skip to >asian_group<
(5) Native Hawaiian
(6) Other Pacific Islander--Skip to >pacific_group<
(7) Some other race--Skip to >othr_race<

D Don't know
R Refused

SPECS: If "D" or "R", store 0 in each variable. Store answers of (1) - (7) as follows:

Current answer	Stored value	Variable
1	1	>RACECB01_6<
2	1	>RACECB02_6<
3	1	>RACECB03_6<
5	1	>RACECB11_6<
7	1	>RACECB15_6<

Store 0 in all variables without a value of 1.

OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at the following locations:
RACECB01_6: location 127
RACECB02_6: location 128
RACECB03_6: location 129
RACECB11_6: location 137
RACECB15_6: location 141

>othr_race<

What is the name of (your/...'s) race?

_____ >othr_race1<
(allow for 19 characters)

D Don't know
R Refused

SPECS: If "D" or "R", store a "Blank."

OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at location 180.

>amer_ind<

What is the name of (your/...'s) enrolled or principal tribe?

(H) HELP

_____ >amer_ind1<
(allow for 19 characters)

D Don't know
R Refused

SPECS: If "D" or "R", store a "Blank."

OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at location 142.

>H_AMERIND<

Add a pop-up or help screen for agents to select the following for the >amer_ind< screen.

Cherokee	**Blackfoot**
Navajo	**Chickasaw**
Chippewa	**Potawatomi**
Sioux	**Tohono O'Odham**
Choctaw	**Pima**
Pueblo	**Tlingit**
Apache	**Seminole**
Iroquois	**Alaskan Athabaskans**
Lumbee	**Cheyenne**
Creek	**Comanche**

SPECS: More than one category is acceptable. When storing more than one category selection, use white space delimiter between the selections.

>asian_group<

To what Asian group (do you/does...) belong? (READ CATEGORIES.)

(1) Asian Indian
(2) Chinese
(3) Filipino
(4) Japanese
(5) Korean
(6) Vietnamese
(7) Other Asian--Skip to >othr_asian<

D Don't know
R Refused

SPECS: If "D" or "R", store 0. Store answers of (1) - (7) as follows:

Current answer	Stored value	Variable
1	1	>RACECB04_6<
2	1	>RACECB05_6<
3	1	>RACECB06_6<
4	1	>RACECB07_6<
5	1	>RACECB08_6<
6	1	>RACECB09_6<
7	1	>RACECB10_6<

OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at the following locations:

RACECB04_6: location 130
RACECB05_6: location 131
RACECB06_6: location 132
RACECB07_6: location 133
RACECB08_6: location 134
RACECB09_6: location 135
RACECB10_6: location 136

>othr_asian<

What other Asian group (do you/does...) belong?

(H) HELP

_____ >othr_asian1<
(allow 9 characters)

D Don't know
R Refused

SPECS: If "D" or "R", store a "Blank."

NOTE: If answers reported for both >othr_asian1< and >othr_pacific1<, combine into one output variable and store in >AISPIWIN_6<. When storing , use white space delimiter between the two answers. Otherwise, store single answer in >AISPIWIN_6<.

OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at location 161.

>H_OTHRASIAN<

Add a pop-up or help screen for agents to select the following for the >othr_asian< screen:

Cambodian
Hmong
Indonesian
Pakistani
Laotian
Thai

SPECS: More than one category is acceptable.

>pacific_group<

SPECS: More than one category is acceptable.

To what Pacific Islander group (do you/does ...) belong? READ CATEGORIES.

(1) Guamanian or Chamorro
(2) Samoan
(3) Other Pacific Islander--Skip to >othr_pacific<

D Don't know
R Refused

SPECS: If "1", store 1 in >RACECB12_6<. If "2", store 1 in >RACECB13_6<. If "3", store 1 in >RACECB14_6<. If "D" or "R", store a "Blank" in >RACECB14_6<.

OUTPUT SPECS: Store in Record 2 for D-1 or D-10, at the following locations:
RACECB12_6: location 138
RACECB13_6: location 139
RACECB14_6: location 140

>othr_pacific<

What other Pacific Islander group (do you/does...) belong?

(H) HELP

_____ >othr_pacific1<
(allow 9 characters)

D Don't know
R Refused

SPECS: If "D" or "R", store a "Blank."

NOTE: If answers reported for both >othr_asian1< and >othr_pacific1<, combine into one output variable and store in >AISPIWIN_6<. When storing, use white space delimiter between the two answers. Otherwise, store one answer in >AISPIWIN_6<.
OUTPUT SPECS: Store in Record 2, for D-1 or D-10, at location 161.

>H_OTHPACIF<

Add a pop-up or help screen for agents to select the following for the >othr_pacific< screen:

Fijian
Palauan
Tahitian
Tongan

SPECS: More than one category is acceptable.

If CEFU case, skip to >THE END< of CEFU script. Otherwise, go to >closing<.

>closing<

SPECS: If custsat=missing, follow path A. If custsat=1, follow path B.

A:
This completes all the questions. Thank you for taking part in Census 2000.

B:
This completes all the questions. Thank you for taking part in Census 2000.

Before you hang-up, we would appreciate feedback regarding the service you received today. I'm going to transfer you to our automated customer satisfaction survey, which on average takes less than 3 minutes to complete.

Programming note: Include the following on this screen.

IF ASKED WHAT TO DO ABOUT THE CENSUS FORM, PLEASE RESPOND—

Since you have given me your census information, you may discard the form you received in the mail.

SPECS: If all of the following fields in >mail_add< are not blank: housestreet1, city1, state1, zip1 and a complete interview is collected, then set Nomail=1.

END INTERVIEW

DEFINITION OF COMPLETED INTERVIEW

TO QUALIFY AS A COMPLETED INTERVIEW--
There must be complete answers* for any two of the following questions for each person on the roster:

>relation< (except person on line 1)

>sex1<

>age<

if >hisp_origin1< = 2 OR if >hisp_origin1< = 1, then >hisp_origin2< must be answered

>race<

*** “Don't know” or “Refused” do not qualify as an “answer.”**

Appendix E

United States Census 2000

Official Internet Form

U.S. Department of Commerce
Bureau of the Census

Start Here!

Instructions:

1. Help is available throughout the form by clicking on the -> Text Links <-. If help links do not function properly, manually open a new/different browser window to: **http://www.2000.census.gov/2k/formhelp.html**
2. Verify this form's authenticity to help protect your information.
3. Use your window scroll bars to move around the form and your screen arrow/pointer to position your text cursor inside boxes for entering text.
4. **DO NOT** use your keyboard's 'Return' / 'Enter' key or your web browser's 'Back' / 'Previous Page' button.
5. Please begin with question 1 below.

1. **How many people were living or staying in this house, apartment, or mobile home on April 1, 2000?**

Number of people

INCLUDE in this number:

- foster children, roomers or housemates
- people staying here on April 1, 2000 who have no other permanent place to stay
- people living here most of the time while working, even if they have another place to live

DO NOT INCLUDE in this number:

- college students living away while attending college
- people in a correctional facility, nursing home, or mental hospital on April 1, 2000
- Armed Forces personnel living somewhere else
- people who live or stay at another place most of the time

2. **Is this house, apartment, or mobile home** -- *Mark ONE box.*

- ☐ Owned by you or someone in this household with a mortgage or loan?
- ☐ Owned by you or someone in this household free and clear (without a mortgage or loan)?
- ☐ Rented for cash rent?
- ☐ Occupied without payment of cash rent?

3. **Please answer the following questions for each person living in this house, apartment or mobile home. Start with the name of one of the people living here who owns, is buying, or rents this house, apartment, or mobile home. If there is no such person, start with any adult living or staying here. We will refer to this person as Person 1.**

What is this person's name? *Enter name below.*

Last Name First Name MI

4. **What is Person 1's telephone number?** *We may call the person if we don't understand an answer.*

Area Code + Number

- -

5. **What is Person 1's sex?** *Mark ONE box.*

☐ Male ☐ Female

6. **What is Person 1's age and what is Person 1's date of birth?**
Enter numbers in boxes.

Age on April 1, 2000 Month Day Year of Birth

NOTE: Please answer BOTH Questions 7 and 8.

7. **Is Person 1 Spanish / Hispanic / Latino?** *Mark the **"NO"** box if **not** Spanish / Hispanic / Latino.*

☐ **No**, not Spanish / Hispanic / Latino ☐ Yes, Puerto Rican
☐ Yes, Mexican, Mexican Am., Chicano ☐ Yes, Cuban
☐ Yes, other Spanish / Hispanic / Latino -- *Enter group below.*

8. **What is Person 1's race?** ***Mark one or more races*** *to indicate what this person considers himself/herself to be.*

☐ White
☐ Black, African Am., or Negro
☐ American Indian or Alaska Native -- *Enter name of enrolled or principal tribe below.*

☐ Asian Indian ☐ Japanese ☐ Native Hawaiian
☐ Chinese ☐ Korean ☐ Guamanian or Chamorro
☐ Filipino ☐ Vietnamese ☐ Samoan
☐ Other Asian -- *Enter race below.* ☐ Other Pacific Islander -- *Enter race below.*

☐ Some other race -- *Enter race below.*

If more people live here, continue with Person 2 below.
If not, go to the 'Sending This Form' section at the bottom of this form.

Person 2

Your answers are important!
Every person in the Census counts.

1. What is Person 2's name? *Enter name below.*

Last Name | First Name | MI

2. How is this person related to Person 1? *Mark ONE box.*

- ☐ Husband/wife
- ☐ Natural-born son/daughter
- ☐ Adopted son/daughter
- ☐ Stepson/stepdaughter
- ☐ Brother/sister
- ☐ Father/mother
- ☐ Grandchild
- ☐ Parent-in-law
- ☐ Son-in-law/daughter-in-law
- ☐ Other relative -- *Enter exact relationship.*

If NOT RELATED to Person 1:

- ☐ Roomer, boarder
- ☐ Housemate, roommate
- ☐ Unmarried partner
- ☐ Foster child
- ☐ Other nonrelative

3. What is this person's sex?

☐ Male ☐ Female

4. What is this person's age and what is this person's date of birth?
Enter numbers in boxes.

Age on April 1, 2000 | Month | Day | Year of Birth

NOTE: Please answer BOTH Questions 5 and 6.

5. **Is this person Spanish / Hispanic / Latino?** *Mark "NO" box if **not** Spanish / Hispanic / Latino.*

- ☐ **No**, not Spanish / Hispanic / Latino
- ☐ Yes, Puerto Rican
- ☐ Yes, Mexican, Mexican Am., Chicano
- ☐ Yes, Cuban
- ☐ Yes, Other Spanish / Hispanic / Latino -- *Enter group below.*

6. **What is this person's race?** ***Mark one or more races*** *to indicate what this person considers himself/herself to be.*

- ☐ White
- ☐ Black, African Am., or Negro
- ☐ American Indian or Alaska Native -- *Enter the name of enrolled or principal tribe below.*
- ☐ Asian Indian
- ☐ Japanese
- ☐ Native Hawaiian
- ☐ Chinese
- ☐ Korean
- ☐ Guamanian or Chamorro
- ☐ Filipino
- ☐ Vietnamese
- ☐ Samoan
- ☐ Other Asian -- *Enter race below.*
- ☐ Other Pacific Islander -- *Enter race below.*
- ☐ Some other race -- *Enter race below.*

If more people live here, continue with Person 3 below.
If not, go to 'Sending This Form' at the bottom of this page.

Person 3

Census information helps your community get financial assistance for roads, hospitals, schools, and more.

1. **What is Person 3's name?** *Enter name below.*

Last Name | First Name | MI

2. **How is this person related to Person 1?** *Mark ONE box.*

- ☐ Husband/wife
- ☐ Natural-born son/daughter
- ☐ Adopted son/daughter
- ☐ Stepson/stepdaughter
- ☐ Brother/sister
- ☐ Father/mother
- ☐ Grandchild

If NOT RELATED to Person 1:

- ☐ Roomer, boarder
- ☐ Housemate, roommate
- ☐ Unmarried partner
- ☐ Foster child
- ☐ Other nonrelative

- ☐ Parent-in-law
- ☐ Son-in-law/daughter-in-law
- ☐ Other relative -- *Enter exact relationship.*

3. What is this person's sex?

☐ Male ☐ Female

4. What is this person's age and what is this person's date of birth?
Enter numbers in boxes.

Age on April 1, 2000	Month	Day	Year of Birth

NOTE: Please answer BOTH Questions 5 and 6.

5. Is this person Spanish / Hispanic / Latino? *Mark **"NO"** box if **not** Spanish / Hispanic / Latino.*

- ☐ **No**, not Spanish / Hispanic / Latino
- ☐ Yes, Puerto Rican
- ☐ Yes, Mexican, Mexican Am., Chicano
- ☐ Yes, Cuban
- ☐ Yes, Other Spanish / Hispanic / Latino -- *Enter group below.*

6. What is this person's race? ***Mark one or more races*** *to indicate what this person considers himself/herself to be.*

- ☐ White
- ☐ Black, African Am., or Negro
- ☐ American Indian or Alaska Native -- *Enter the name of enrolled or principal tribe below.*
- ☐ Asian Indian
- ☐ Japanese
- ☐ Native Hawaiian
- ☐ Chinese
- ☐ Korean
- ☐ Guamanian or Chamorro
- ☐ Filipino
- ☐ Vietnamese
- ☐ Samoan
- ☐ Other Asian -- *Enter race below.*
- ☐ Other Pacific Islander -- *Enter race below.*
- ☐ Some other race -- *Enter race below.*

If more people live here, continue with Person 4 below.
If not, go to 'Sending This Form' at the bottom of this page.

Person 4

Information about children helps your community plan for child care, education, and recreation.

1. **What is Person 4's name?** *Enter name below.*

Last Name | First Name | MI

2. **How is this person related to Person 1?** *Mark ONE box.*

- ☐ Husband/wife
- ☐ Natural-born son/daughter
- ☐ Adopted son/daughter
- ☐ Stepson/stepdaughter
- ☐ Brother/sister
- ☐ Father/mother
- ☐ Grandchild
- ☐ Parent-in-law
- ☐ Son-in-law/daughter-in-law
- ☐ Other relative -- *Enter exact relationship.*

If NOT RELATED to Person 1:

- ☐ Roomer, boarder
- ☐ Housemate, roommate
- ☐ Unmarried partner
- ☐ Foster child
- ☐ Other nonrelative

3. **What is this person's sex?**

☐ Male ☐ Female

4. **What is this person's age and what is this person's date of birth?**
Enter numbers in boxes.

Age on April 1, 2000 | Month | Day | Year of Birth

NOTE: Please answer BOTH Questions 5 and 6.

5. **Is this person Spanish / Hispanic / Latino?** *Mark "NO" box if **not** Spanish / Hispanic / Latino.*

- ☐ **No**, not Spanish / Hispanic / Latino
- ☐ Yes, Mexican, Mexican Am., Chicano
- ☐ Yes, Puerto Rican
- ☐ Yes, Cuban
- ☐ Yes, Other Spanish / Hispanic / Latino -- *Enter group below.*

6. **What is this person's race?** ***Mark one or more races*** *to indicate what this person considers himself/herself to be.*

- ☐ White
- ☐ Black, African Am., or Negro
- ☐ American Indian or Alaska Native -- *Enter the name of enrolled or principal tribe below.*

- ☐ Asian Indian
- ☐ Japanese
- ☐ Native Hawaiian
- ☐ Chinese
- ☐ Korean
- ☐ Guamanian or Chamorro
- ☐ Filipino
- ☐ Vietnamese
- ☐ Samoan
- ☐ Other Asian -- *Enter race below.*
- ☐ Other Pacific Islander -- *Enter race below.*

- ☐ Some other race -- *Enter race below.*

If more people live here, continue with Person 5 below.
If not, go to 'Sending This Form' at the bottom of this page.

Person 5

Knowing about age, race, and sex helps your community better meet the needs of everyone.

1. **What is Person 5's name?** *Enter name below.*

Last Name | First Name | MI

2. **How is this person related to Person 1?** *Mark ONE box.*

- ☐ Husband/wife
- ☐ Natural-born son/daughter
- ☐ Adopted son/daughter
- ☐ Stepson/stepdaughter
- ☐ Brother/sister
- ☐ Father/mother
- ☐ Grandchild
- ☐ Parent-in-law
- ☐ Son-in-law/daughter-in-law
- ☐ Other relative -- *Enter exact relationship.*

If NOT RELATED to Person 1:

- ☐ Roomer, boarder
- ☐ Housemate, roommate
- ☐ Unmarried partner
- ☐ Foster child
- ☐ Other nonrelative

3. **What is this person's sex?**

- ☐ Male
- ☐ Female

4. **What is this person's age and what is this person's date of birth?**
Enter numbers in boxes.

Age on April 1, 2000	Month	Day	Year of Birth

NOTE: Please answer BOTH Questions 5 and 6.

5. **Is this person Spanish / Hispanic / Latino?** *Mark "NO" box if **not** Spanish / Hispanic / Latino.*

- ☐ **No**, not Spanish / Hispanic / Latino
- ☐ Yes, Puerto Rican
- ☐ Yes, Mexican, Mexican Am., Chicano
- ☐ Yes, Cuban
- ☐ Yes, Other Spanish / Hispanic / Latino -- *Enter group below.*

6. **What is this person's race?** ***Mark one or more races*** *to indicate what this person considers himself/herself to be.*

- ☐ White
- ☐ Black, African Am., or Negro
- ☐ American Indian or Alaska Native -- *Enter the name of enrolled or principal tribe below.*
- ☐ Asian Indian
- ☐ Japanese
- ☐ Native Hawaiian
- ☐ Chinese
- ☐ Korean
- ☐ Guamanian or Chamorro
- ☐ Filipino
- ☐ Vietnamese
- ☐ Samoan
- ☐ Other Asian -- *Enter race below.*
- ☐ Other Pacific Islander -- *Enter race below.*
- ☐ Some other race -- *Enter race below.*

If more people live here, continue with Person 6 below.
If not, go to 'Sending This Form' at the bottom of this page.

Person 6

Your answers help your community plan for the future.

1. **What is Person 6's name?** *Enter name below.*

Last Name	First Name	MI

2. **How is this person related to Person 1?** *Mark ONE box.*

- ☐ Husband/wife
- ☐ Natural-born son/daughter
- ☐ Adopted son/daughter
- ☐ Stepson/stepdaughter
- ☐ Brother/sister
- ☐ Father/mother
- ☐ Grandchild
- ☐ Parent-in-law
- ☐ Son-in-law/daughter-in-law
- ☐ Other relative -- *Enter exact relationship.* ☐

If NOT RELATED to Person 1:

- ☐ Roomer, boarder
- ☐ Housemate, roommate
- ☐ Unmarried partner
- ☐ Foster child
- ☐ Other nonrelative

3. **What is this person's sex?**

☐ Male ☐ Female

4. **What is this person's age and what is this person's date of birth?**
Enter numbers in boxes.

Age on April 1, 2000	Month	Day	Year of Birth
☐	☐	☐	☐

NOTE: Please answer BOTH Questions 5 and 6.

5. **Is this person Spanish / Hispanic / Latino?** *Mark "NO" box if **not** Spanish / Hispanic / Latino.*

- ☐ **No**, not Spanish / Hispanic / Latino
- ☐ Yes, Mexican, Mexican Am., Chicano
- ☐ Yes, Other Spanish / Hispanic / Latino -- *Enter group below.* ☐
- ☐ Yes, Puerto Rican
- ☐ Yes, Cuban

6. **What is this person's race?** ***Mark one or more races*** *to indicate what this person considers himself/herself to be.*

- ☐ White
- ☐ Black, African Am., or Negro
- ☐ American Indian or Alaska Native -- *Enter the name of enrolled or principal tribe below.* ☐
- ☐ Asian Indian
- ☐ Japanese
- ☐ Native Hawaiian
- ☐ Chinese
- ☐ Korean
- ☐ Guamanian or Chamorro

☐ Filipino ☐ Vietnamese ☐ Samoan

☐ Other Asian -- *Enter race below.* ☐ Other Pacific Islander -- *Enter race below.*

☐ Some other race -- *Enter race below.*

If more people live here, continue with Person 7 below.
If not, go to 'Sending This Form' at the bottom of this page.

Persons 7-12

If you didn't have room *(above)* **to list everyone who lives in this house or apartment, please list the others below.** *You may be contacted by the Census Bureau for more information about these people.*

	Last Name	First Name	MI
Person 7			
Person 8			
Person 9			
Person 10			
Person 11			
Person 12			

- End of Census Form -

OMB No.: 0607-0856
Approval Expires 12/31/2000

Sending This Form

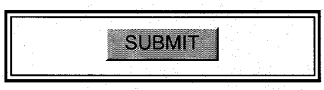
If you have finished this form:

- Before sending this form, please go back to the top of this form and review your information now for completeness and accuracy.

  Did you:
  1. Answer the questions for all persons in your household?
  2. Review your entries for completeness and accuracy?
  3. Include a telephone number for Person 1?

  You will have only one opportunity to send this form.

- If you are satisfied you have entered all information completely and correctly, please send the form back to the Census Bureau by using the 'SUBMIT' button below. A 'confirmation' page will be sent to you as soon as your information is received.
- DO NOT attempt to re-send this form once you receive a confirmation page.
- If you do not receive a confirmation page, please complete your paper form and return it in the pre-addressed envelope.



Appendix F

Transcriptions of Responses to the Open-ended User Satisfaction Questions

Please tell us what you disliked about the computerized questionnaire after the beep.

1 You could have made it all touchtone except for the name.
2 No comment.
3 I disliked the fact the voice recognition software is inadequate-it cuts you off when you're trying to speak.
4 It was too slow in repeating the information.
5 It's basically inane.
6 System's very very slow much too slow to operate efficiently would have been much easier just to fill out the questionnaire.
7 It was too long.
8 Too slow.
9 Like to be able to go back & correct a question if possible, to be able to back up.
10 Takes way too much time I could do it much much faster by filling it out and mailing it in.
11 It took me a long time to respond in this system I felt like I was going to fall asleep toward the end of the phone call.
12 Redundant; you ask the date of birth, then ask you the age, you ask the relationship; daughter/son, then you ask if male or female.
13 Had to spell everybody's name what is this race thing we are all Americans.
14 You should follow the format on the printed form - when summarize you leave information out.
15 Redundant. I mailed you the form.
16 Too slow.
17 It takes way too long, you repeat yourself way too much, after answering one person's questions I quite remember everything else, thank you.
18 There wasn't anything I disliked.
19 Providing the answers was very slow & time consuming much quicker to complete the written form.
20 Asked to repeat too many times.
21 Too slow.
22 Nothing.
23 The information on the names of the other members takes too long.
24 Too slow.
25 It required too much time.
26 Takes forever.

Please tell us what you liked about the computerized questionnaire after the beep.

1 It was simple-I understood everything-it was just an easy way to do it.
2 Easy.
3 Convenient and fast.
4 The ease in use.
5 Very efficient.
6 Well I thought it was very easy but it could have gone a little faster but it was convenient.
7 The calling card.
8 Easier.
9.
10 Easy to use.
11 It was quick & easy, succinct & to the point.
12 It was convenient.
13.
14 Didn't have to mail anything.
15 The questions were very clear-didn't take too long.
16 Just a lot easier to do it over the phone then to send it in.
17 Free calling card.
18 The questions were direct and easy to answer. not much invasion of privacy.
19 I think it was very clear and easy to understand.
20 Simple.
21 No I just like being able to do it over the phone here.
22 Its convenient.
23 It's easy.
24 Short easy to do no mailing.
25 Easy to answer questions and understand.
26 Well I don't have to fill out this form and mail it in.
27 Clear, concise, direct, simple.
28 Clear and to the point.
29 The convenience of the questionnaire.

30 Convenient.
31 It was slow.
32 It was very easy to follow and understand and easy to be able to call and get the census done.
33 Well I liked the computerized questionnaire because it was fast and easy-the incentive of having the 30-minute calling card actually got me to turn it in-uh to call in-i.
34 Convenience.
35
36 It took only 10 minutes.
37 Easy.
38 Prepaid phone card.
39 It was easy and convenient to use.
40 Convenience, knowing the thing was done.
41 It was short and easy.
42 The ease of the system.
43 Easy.
44 It was quick.
45 It's quick and easy.
46 Its easy its simple.
47 Much easier than having to mail it in.
48 Because it's simple and quick.
49 Simple.
50 I like talking more convenient.
51 Because its simple.
52 Clear and to the point.
53 Like getting the free calling card.
54
55 It was very precise and short..
56 Convenience.
57
58 Easy & fun, convenient.
59
60 It's easy.
61 It was straightforward clear and simple.
62 It's fast and easy, and it doesn't require a lot of effort, you can just get it done.
63 It was very clear to hear understand and use.
64
65 Clear and concise.
66 Very concise.
67 Easy-to-use.
68 It's a, mailing the, the uhh, form.
69 Very simple straight forward to the point.
70 It's a little slow but it's okay.
71 It was simple.
72 It was easy.
73 Save time going to the post office.
74
75 Ease of use.
76 It's quick, easy and there was no wait on the phone.
77 Short and to the point.
78 Voice activated.
79 Didn't take too long.
80 That I didn't have to mail it in that I could just call over the phone and the thirty minute calling card was very inviting.
81 Easy to use.
82 Very convenient/like calling card.
83 It s quick.
84 Its very easy.
85 It was quick and efficient.
86 Very comprehensive & clear especially when its repeated in your own voice.
87 Pretty easy to follow.
88 The lady's voice.
89 It made it easier not having to worry about getting something to the mailbox.
90
91 Its fast and easy to understand.
92 You could call in 24 hours.
93 It was easy.
94
95 Simple.
96 Quick.
97 Quick.
98 It was slow and concise.
99

100 Everything was very clear about it and I liked getting the voice on the voice on the telephone and I thought it went very well yes I liked it.
101 No.
102 Calling card.
103 It was very clear and straight forward it was easy to understand.
104
105 It was convenient and didn't take too much time and it enabled me to answer the census after I lost my mailed form.
106 It's convenient.
107 It's uh easy to use and I don't have to sit down and answer all the questions by reading them and Then figuring them out and it's just a better system.
108 Brief, clear.
109 It eliminates having to fill out a form and mail it in.
110 Short and a calling card.
111 Quick, easy and I got a phone card.
112 Easy to use.
113 It was easy to use-no hassles-and I got it done quickly.
114 Easy to do at any time.
115 Simple efficient easy to understand and not too much of my own time.
116
117 Not someone knocking at the front door.
118 Its just easier to call and pass on information and not have to fool around with that meddling bag.
119 Quick, easy, to the point, you have everything taken care of with the phone call, you didn't have to mail anything out.
120 Convenience.
121 I didn't have to fill in & send in form & gets it done all at once.
122 I didn't have to fill out this long piece of paper and send it off.
123
124 It was simple, clear and concise.
125 I like the free phone card.
126 Easy.
127 Very direct and very easy to follow.
128
129 Like knowing she has it done and that she has the 30 minutes of free phone time..
130 this is a little bit to fast right now I uh I cannot answer right now it was ok I liked it thank you.
131 Easy questions.
132 Convenient.
133 It was nice not having to type in a bunch of stuff.
134 It was easy to understand.

Please tell us what you did not understand after the beep.

1 To begin with I was talking before the people.
2 You ask some unnecessary information.
3
4 I understood everything you asked.
5 Nothing.
6 The flow to the questions was too abrupt.
7 I felt the system was a little time consuming.
8 Relationships.

Please tell us what the computer did not understand after the beep.

1
2 Understood everything.
3 About 2/3 of what I said.
4 When I gave my origin, it misunderstood.
5 It repeated the first series of information about person number 1 twice.
6 I'm not sure.
7 I had to repeat several times certain answers.
8 Some of my yes and no answers.
9 Get a life.
10 Did not give enough time for birthdate.
11 When you playback the info to see if it is correct you've cuttin' off part of a last name.
12 It didn't understand wife and I had to spell out my relationship w-I-f-e.
13 I'm not understanding what the computer is wanting.
14 Did not give me an opportunity to go back and correct my husbands name.
15 Computer understood everything.
16 Ask me to spell my relationships.
17 It did not allow enough time to answer the questions on birth, birthdates.
18 Recording of our names is unclear.
19 The computer understood everything I said.
20 I wasn't able to say my whole birthdate, it cut me off, and when I tried to correct it, it still did not give me time..
21 The computer did not take the birth date of my husband the first time I put it in.

22
23 Other nonrelative.
24 Number 4.

Please tell us what was confusing or frustrating after the beep.

1 It kept dragging on when I had already answered questions.
2 Redundant.
3 In the areas of origin and race.
4 I would prefer not to have to end put numbers on the key pad, it would be better if everything could be spoken.
5 Unable to correct, incorrect answer.
6 Item 2.
7 The fact is I filled out a form and sent it in over a month ago.
8 Why do you ask to spell the relationship of child I do not understand that.
9 You had asked what they were natural born son or daughter then you turn around and ask me what their sex is-well if they're a son obviously they're male-if they're a daughter.
10 Race.
11 Having to constantly repeat everything in slow motion.
12 It was frustrating that the computer was so slow.
13 Too long.
14 For my stepchildren we have joint custody and the questionnaire did not make it clear to which parent should claim the child - children.
15 Too slow.
16 Way too slow.
17 The operator spoke to slowly and it took too long to go through the questionnaire.
18
19 Determining race.
20 What exactly do you mean by spell the name of how the person is related to you, for an example such as my natural born son or daughter, you want me to actually spell out.
21 Some things are personal and not anyone's business it asks some questions that were not what I would call should be given out to public.
22 What was confusing at the very beginning was #9, question #9 so I think I answered mine incorrectly.
23 Takes too long.
24 Frustrating is that...why do you need to know the person's date of birth. what's the sense of it? this is just too personal...that's my point..
25 Question number 1 was confusing.
26 Too hard to make changes.
27 Were the person race and origins.
28 The questions take too much time and repetition.
29.

Please tell us your suggestions about improving the computerized questionnaire after the beep.

1 Everything was fine.
2 End the questionnaire now.
3 For family you should be able to enter info for like last name and race and all that for the whole family at once instead of having to do it individually.
4
5
6
7 I have none it is very satisfactory.
8 I can't improve upon it. I think it's nicely done..
9 None.
10 Fixing the voice recognition software-allowing more time to answer the questions.
11 All voice input would be preferable.
12 Don't have any.
13
14 None.
15
16 No suggestions it's fine.
17 Don't ask for suggestions or how to rate this thing.
18 No suggestions at this time.
19 Just like I said before it could have gone a little faster maybe but overall its pretty good.
20 No suggestions everything worked good.
21 Frustrated she didn't get peter j. in asq confirmation.
22 None.
23 I haven't any.
24 No improvement, fine.
25 Goes a little slow.
26
27 Um the only thing I can think of is in the race thing my last name is Alvarado but my mother was Irish so I'm half Irish & half Spanish so if there was something to figure.

28 Ohh just try to make it be quicker.
29 Hum you could send numbers for each letter of the alphabet and then people could input their names with the numeric keypad on their telephone.
30 No suggestions.
31 The speech can be more quickly.
32 Tie in the same-the data-so if they're parents you know automatically that they're gonna be the same last name or at least the sexes -you know-if they're sexes could be tied.
33 No suggestions.
34 None.
35 Oh I don't see any need for improvement it's a hell of a lot better then doing your taxes.
36
37 None.
38 None at this time.
39 Don't have any.
40
41 I have no suggestions.
42 Really don't have any.
43 No suggestions.
44 When it come to the house first id ask if you own your house and then id say is it paid for or do you have a mortgage or id make it two questions it was a fine survey as.
45 I have none.
46 We should be able to answer before after the beep comes on.
47 None.
48 I have no suggestion its just fine.
49 Do not have any.
50 I think you should send the computerized questionnaire and not bother sending the long questionnaire.
51 No suggestions.
52 You'd have to ask Microsoft and Bill Gates I understand they do a lot of government work.
53 I have none it worked very well.
54 Don't have any, its fine.
55 Being able to key more in instead of speaking it.
56 If you missed a question or put in the wrong answer you should be able to go back and correct it.
57
58 Too much time - took 20 minutes - a little inconvenient.
59 People should have option as to whether or not they want to answer these questions.
60 No suggestions-it was very efficient-thank you very much.
61 Well I don't have really any suggestions-I think it was very well put together and just very effective.
62 No suggestions.
63 None.
64 No suggestions.
65 Spelling some of easy things such as natural born son is cumbersome that could have gone more quickly.
66 Using numbers instead of yes and no would have been helpful.
67 None.
68 Make this entire process available via the website much easier to type this information then to speak it to the phone and to spell each of the names and so on and so fort.
69 I don't have any.
70 Make it faster.
71 No suggestions. you did fine computer.
72 I don't have any suggestions.
73 None.
74 I have none.
75 There is none.
76
77 Less questions.
78 Leave it the way it is.
79 None.
80 I have none.
81 I don't have any.
82 Have a real person do the questioning.
83 None.
84 I liked it just the way it was.
85 It's fine like it is.
86 Right now I don't have any.
87 I'd like to verify my pin number on the calling card..
88 I have no comment.
89 None.
90 Don't need any.
91
92
93 No.
94 I have no suggestions.
95 Seems to be perfect..

96 No suggestions.
97 Don't really have any.
98 I have none.
99 I have none.
100 Lets make it faster.
101 I have none.
102 I really have no suggestions I think it was just fine.
103 I don't have any.
104 Fine the way it is.
105 Just keep improving it thank you.
106 None.
107 I don't know if you can because it takes too much time to do it by the computer compared to the Time to do it on paper.
108 Less pause after answering questions.
109 Allow more speaking time.
110 No suggestions for improvement.
111 Don't have any. continue to march man.
112 None.
113 None.
114 No suggestions.
115 No suggestions.
116 I think everything was fine it should stay as it is.
117 I don't have any.
118 None.
119 None.
120 I don't have any.
121 No suggestions.
122 The only thing that was slightly confusing was that on the written, umm census it had after number three it had last name and then first name and when you asked me, the co.
123 Eliminate redundant questions.
124 Have the operator speak faster.
125
126 None.
127 It's not improving the questionnaire.
128 Follow the form on the questionnaire instead of bouncing around.
129 No clue.
130 Cut down on the wait time in between.
131 Check your mail first.
132 Keep up the good work.
133 I don't have any.
134 Go a little faster.
135 I have none it was fine.
136 None really.
137 None.
138 Have no suggestion.
139 I have none.
140 No suggestions.
141 I have none.
142 None to report.
143 No suggestions.
144 I don't have any I was very satisfied.
145 No suggestions.
146
147 Eliminate the question about race.
148 No suggestions.
149 Everything's fine.
150 No suggestion you're doing fine.
151 Once the directions have been gone over with the first person on our list the people after that you shouldn't have to give the directions over and over so that we have to.
152 Need calling card number before starting the survey.
153 Sounds good to her.
154 Have none.
155 None.
156 None, it was fine.
157
158 An option to answer by touch tone numbers equaling answers.
159 Perhaps somewhere on the form you could inform people how long the questionnaire is going to take, you could give a little bit more time for the people to respond.
160
161
162
163 I don't have any suggestions it worked well.
164 More questions than you had noted on card next time be honest.

165 None.
166 I don't have any suggestions I think everything was fine.
167 None.
168 None.
169
170 No suggestions.
171 It's fine.
172 Need a backup option to correct errors.
173 I don't have any suggestions.
174 Not sure it can be improved, it needs to take so much time ,I just found it to take too much time, that's all.
175 Sounded great to me.
176 None.
177 No suggestions.
178 I can't think of any at this time.
179 Just shorten the id number we only have two hundred and sixty million people I don't know why we need a number 22 characters long.
180 Computerized questionnaire seems fine, don't like the automated phone for correcting mistakes.
181 I think it's perfect the way it is.
182 Possibly people could determine whether to questions orally or typing in questions using buttons on the telephone or e-mail would be faster.
183 You should tell them that you are going to ask them to spell name afterwards you are going to repeat the information twice if you hear the little thing then you know you.
184 Its too long need to shorten it up a little bit.
185 It's ok the way it is..
186 Provide an option to press a key and have the questionnaire pull it while you look at the form.
187 It's easy, but just don't ask so much personal questions...I don't think it's all that necessary. how many people live in the house and where you live that you.
188 Worked fine for me nothing wrong.
189 Speed it up.
190 None.
191 None.
192 Offer the form on-line via the internet.
193 None.
194 I don't know anything to tell you-i think its great-I had no problems.
195 Move a little bit faster.
196 None.
197 None.
198 Read back spellings of all names.
199 Very clear.
200 I don't have any.
201 Put the id number where it can be located more easily and also indicate if this is a free call.
202 I believe its ok just like it is.
203 Answers could be done by pressing number instead of having to list the answer.
204 None for me, everything was fine.
205 None.
206
207 None.
208 None given.
209 No suggestions.
210 Ask if members in house have the same last name instead of going through the whole procedure.
211 I don't have any suggestions for you.
212 No suggestions.
213 More time for some of the questions..
214 No comment on it.
215
216 Speed it up.
217 No suggestions.
218 None.
219 I'm surprised that occupation is not asked.
220 Would have numbers for each of the answers so could just press the button rather than saying yes Or no and waiting for the pauses.
221 this is first time I have been on anything like this really don't know but it is fine better than having to push all these buttons to make a telephone call push if you want.
222
223 Make it faster.
224 None I thought it was great.
225 Not to make it so long.
226 Make it online accessible.
227 None.
228
229 To move along more quickly.
230 None.
231 No suggestions.

232 Uh it it its ok.
233 Speed it up.
234 I have no suggestions for that.
235 None really.
236 No need fine.
237 None.
238
239
240 I have none at this point.

Appendix G

Tables corresponding to figures in the text.

Figure 1. Results: All Households

	No Incentive			Incentive		
	CATI	**ASQ**	**Internet**	**CATI**	**ASQ**	**Internet**
Mail	**1689**	**1521**	**1628**	**1292**	**1234**	**1343**
Cati	**34**	**1**	**7**	**422**	**17**	**8**
ASQ	**0**	**17**	**0**	**0**	**437**	**0**
Internet	**1**	**0**	**93**	**0**	**0**	**370**
UAA	**275**	**160**	**290**	**235**	**239**	**219**
No response	**622**	**922**	**609**	**673**	**696**	**684**
Total	**2621**	**2621**	**2627**	**2622**	**2623**	**2624**

Figure 2. Results: High Coverage Area Households

	No Incentive			Incentive		
	CATI	**ASQ**	**Internet**	**CATI**	**ASQ**	**Internet**
Mail	**1440**	**1268**	**1359**	**1089**	**1021**	**1109**
Cati	**28**	**1**	**6**	**374**	**11**	**7**
ASQ	**0**	**12**	**0**	**0**	**383**	**0**
Internet	**1**	**0**	**80**	**0**	**0**	**336**
UAA	**430**	**721**	**433**	**478**	**508**	**500**
No response	**214**	**114**	**237**	**173**	**191**	**165**
Total	**2113**	**2116**	**2115**	**2114**	**2114**	**2624**

Figure 3. Results: Low Coverage Area Households

	No Incentive			Incentive		
	CATI	**ASQ**	**Internet**	**CATI**	**ASQ**	**Internet**
Mail	**249**	**253**	**269**	**203**	**213**	**234**
Cati	**6**	**0**	**1**	**48**	**6**	**1**
ASQ	**0**	**4**	**0**	**0**	**54**	**0**
Internet	**0**	**0**	**13**	**0**	**0**	**34**
UAA	**192**	**201**	**176**	**194**	**190**	**184**
No response	**61**	**46**	**52**	**62**	**48**	**53**
Total	**508**	**504**	**511**	**507**	**511**	**506**

Figure 5. Calling card usage

Value remaining ($)	Number of cards
0 (fully used)	**239**
0.01 to 1.00	**119**
1.01 to 2.00	**43**
2.01 to 3.00	**39**
3.01 to 4.00	**24**
4.01 to 5.00	**31**
5.01 to 6.00	**27**
6.01 to 7.00	**32**
7.01 to 7.49	**16**
7.50 (never used)	**292**

Figure 6. Confidence levels returned by the speech recognizer and number of utterances.

Confidence	Number
26	1
50	2
56	1
59	4
60	9
61	3
62	3
63	6
64	4
65	5
66	4
67	9
68	6
69	10
70	18
71	23
72	26
73	26
74	33
75	19
76	41
77	38
78	53
79	46
80	71
81	70
82	70
83	79
84	114
85	111
86	117
87	160
88	179
89	196
90	250
91	247
92	269
93	351
94	401
95	471
96	587
97	660
98	642
99	58
100	167