

Design of Coverage Probes for Coverage Edit Followup for Census 2000

American Statistical Association, Atlanta, GA, August 2001

Dave Sheppard, U.S. Census Bureau, Decennial Statistical Studies Division

This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

Key Words CATI; questionnaire design; browser

Abstract During Census 2000, a coverage edit was conducted on all forms mailed in to a Data Capture Center or were submitted over the Internet. Two kinds of cases failed the edit - forms from households with more than six persons and forms on which the number of persons for whom data was captured was not equal to the number reported by the respondent as the household size. The coverage edit instrument was a computer assisted telephone interview (CATI) developed in HTML without any personal visits. We used a series of coverage probes to ensure that each household roster was correct before collecting person data for any additional persons. This paper will talk about the design of these probes as well as some preliminary results from the operation.

Why do we conduct Coverage Edit Followup?

When we did the Census last year, most respondents mailed back their form to us. Sometimes, what they wrote confused us. For example, a respondent may have indicated there were three household members, but only entered names and demographic information about two persons. Which is it? To find out, we followed up with respondents by phone - but we had to decide what to ask them. The purpose of this paper is to explain how we decided what to ask them.

The objectives for the 22nd Decennial Census included making unprecedented efforts to count every person and household while also avoiding erroneous enumerations. In simpler words, don't miss people and don't count them where they don't belong. Coverage edit followup was designed to help meet both of these objectives.

By coverage, I'm talking about counting everyone at the correct place. Our residence rules define where each person should be counted. By edit, I mean selecting cases based on a certain criteria. The criteria we set included the types of forms and the way we determine whether a case gets selected.

Coverage Edits for Census 2000

In Census 2000, a computer based coverage edit operation was conducted to improve the coverage of persons in housing units. The eligible universe included all mail-return (Short form, Long form and Be Counted¹ forms) and Internet questionnaires. Two kinds of cases failed the edit - count discrepancies and large household cases.

Count discrepancy cases result when the respondent reported household size varies from the total number of persons for whom Census information was provided on the form.

Large household cases result when there is a household person count of 7 or more, since the questionnaire only has space to report population data from a maximum of six persons (for Be Counted forms, there is only room for 5 persons, so a household size of 6 or more qualified as a large household case).

Similarly, questionnaires completed by enumerators during the nonresponse followup operation were reviewed to ensure that the data were collected for all persons reported living at the address. Cases that failed the review were revisited or corrected while large household cases had person data collected for all persons at the original time of interview using continuation forms. No telephone followup was necessary for these cases and they are not discussed in this paper.

The Census 2000 coverage edits operation was designed to be done solely by telephone without any personal visits. Census staff designed the coverage edit followup questionnaire and as well as selected the universe of cases for this operation. A subcontractor, Electronic Data Systems (EDS) was contracted to implement the coverage edit operation in conjunction with the Telephone questionnaire assistance program (TQA). The interview was conducted using a computer assisted telephone interview (CATI) instrument which was developed in HTML. Cases eligible for coverage

¹Be Counted Questionnaires were developed for people who did not receive a questionnaire or believed they were not included elsewhere on a census form. Respondents obtained a form from a variety of public locations and mailed the completed form in to a data capture center.

edits were selected from forms processed by June 8, 2000 and these cases were contacted by telephone between May 8th and August 14th, 2000.

Research and Experience

Each time a census is designed, residence rules and rostering methods are refined in attempts to make it more likely respondents will understand the rules and instructions well enough to report the appropriate persons as residents in their household. Additionally, we considered the mode of interview, as well as the respondent burden due to the length.

To accomplish this, we looked at what had been used by past census operations and as well as rostering studies. When designing the Coverage Edit Followup instrument for Census 2000, we considered how best to probe to get a complete household roster. Much research has been done in the past on the best way to go about this including:

- 1990 Census Coverage Edits operation
- Coverage Sampling Survey (1990)
- Living Situation Survey (1994)
- Census 2000 Dress Rehearsal (1998)

The **1990 Census Coverage Edit Followup operation** was conducted over the telephone out of the local offices and the processing centers. The universe eligible for coverage edits included mailback and enumerator forms in its eligible universe and in some instances if the telephone followup was unsuccessful, coverage edit followup was conducted by a personal visit.

Unfortunately, the volume and results of the 1990 Census Coverage Edits followup were not documented. Also, coverage probes were not scripted; rather, clerks marked problem areas on the questionnaires and staff discussed corrections with the respondents when contacted by telephone.

During the 1990 Decennial Census, the Census Bureau conducted the **Coverage Sampling Research**. This study was done with probes which were similar to the list of who to include and who not to include which was listed on each 1990 Census form. Eight questions were asked:

1. Have I missed anyone who usually lives here but might have been away on April 1st, for example staying with friends or relatives, on vacation, on a business trip or in a hospital?
2. Have I missed any small children, foster children, or infants born on or before April 1st ?
3. Have I missed any lodgers, boarders or persons you employ who stayed here on April 1st ?
4. Have I missed anyone who stayed here on April 1st because they have no other home?

5. Referring back to the persons listed on your Census questionnaire, were any of these persons away in the military on April 1st ?
6. Did any of these persons live someplace else while attending a college or university on April 1st ?
7. Were any of these persons in an institution such as a prison, mental hospital, or nursing home on April 1st?
8. Did any of these persons live someplace else but stay here temporarily on April 1st ?

The results from this research indicated that the questions about college students and regular household members who may have been temporarily away confused the respondents. For the telephone sample, both of these questions were responded to positively by close to two percent of a sample of respondents. None of the other six probes resulted in more than 1% of the telephone respondents adding or deleting a person from the household roster they had initially provided.

The **Living Situation Survey** (1994) was done to study the rostering process. It was designed to learn more about household composition, mobility, residency patterns, and the attachment of some people to multiple households. This research addressed the increasingly varied and complex living situations of the population. Results show that additional probe questions can be effective in picking up persons omitted in response to the initial roster question.

Cognitive research was also conducted to gain a better understanding of how respondents understand terms used in census questions. Some of the words studied were "live", "stay", "usual residence", and "household". Results show that neither "usual residence" nor "household" are commonly used among respondents. It was recommended that wording the initial roster question by asking for those persons "who live here permanently or who were staying here temporarily..." could improve respondents' understanding of the rostering task.

The **Census 2000 Dress Rehearsal Coverage Edit Followup** was conducted in 1998 for forms on which the count of names listed on the roster differed from the number of person records that could be data captured. The telephone interview used a paper questionnaire without specific probes read relating to residence rules. The respondent contacted by phone was read the roster on the form and asked if it was correct. If so, then the person count was corrected. If it was not, the roster was corrected to match the person count.

It is important to note that this was not a true dress rehearsal of the coverage edit followup design, since the methodology and technology used for Census 2000 coverage edits operation itself was very different than in dress rehearsal. The design had never been tested prior to Census 2000.

Designing Probes for Coverage Edit Followup in Census 2000 (CEFU instrument design)

An interdivisional team created the CEFU instrument. Many issues were considered. We needed to be very specific in our scripts to ensure that the residence rules would be interpreted correctly. However, we also needed to be mindful that increased respondent burden could lead to a high respondent refusal rate.

The instrument had four main components. The first established that the correct household had been reached. The second reviewed the roster with the respondent and gave the respondent an initial opportunity to correct the roster. Next followed a series of coverage probes to ensure that each person on the household roster was correctly included. These probes in this section were designed around the residence rules and allow the respondent to add or delete names from what was reported on their census form in order to more accurately reflect their household on Census Day. Finally, the instrument collected personal information necessary to create an enumeration record for any persons added to the roster during the roster review or coverage probes or for any persons originally listed on the form for whom there was no room on the form to report.

As in the 1990 census, coverage probes aligned closely with the instructions on who to include which were listed on each Census 2000 form.

There were five specific questions which were designed to add persons who the respondent had not originally included on their form but should have been according to the census residence rules. These were:

Question 1

We've found that sometimes people were left off the census form, especially people who stayed somewhere temporarily or were new to the household, among other reasons. I am going to ask you some questions just to make sure that no one was missed.

Other than those persons you included, were there any children who were living or staying there around the beginning of April? Be sure to consider any newborns, foster children, step children, or children in shared custody arrangements.

Question 2

Other than those you included, were there any relatives, such as aunts, uncles, grandparents, cousins, or any other kinds of relatives who were living or staying there around the beginning of April?

Question 3

Other than those you included, were there any other persons not related to you who were living or staying there around the beginning of April? For example, someone who rents a room from you or a friend staying with you temporarily while looking for a place to live.

Question 4

Other than those you included, were there any persons who were either away temporarily or moving around the beginning of April? For example, a household member who was visiting with friends or relatives, on vacation, on a business trip, or in the process of moving.

Question 5

Think back to the beginning of April. Were there any people staying there who had no other permanent place to stay? Please tell me their names even if you do not consider them to be regular members of your household.

There were four specific questions which were designed to delete persons who the respondent included on the roster but should not have been according to the census residence rules. These were:

Question 6

OK. Now I have some questions about people who were included on your form and might have been counted elsewhere. The census counts people in places where groups of people stay such as colleges, nursing homes, jails and the like. For these next questions, please think about the people who were included on your form.

Were any of these people college students living or staying away around the beginning of April?

Question 7

Were any of these people members of the U.S. Armed Forces living or staying away around the beginning of April?

Question 8

Were any of these people living away in a place such as a nursing home, prison or jail, mental hospital, or similar place on April 1st?

Question 9

Some people have more than one place to live. Examples include a second residence where they stay to be closer to work, a friend's or relative's home, or a vacation home. Did any of the people on the list I read you earlier have another place where they live or stay?

Note that the structure of most of the probes is consistent. The time frame referenced is "around the beginning of April". Also, examples are given with in

most questions. Add questions began with “Other than those persons you included,...” and the delete probes began with “Were any of these people...”. We felt that by having a consistent structure, it would make our probes more understandable to the respondents.

There was also an opportunity for the respondent to indicate that there was more than one person included on the roster who represented the same person. Additionally, the instrument allowed the respondent to interrupt the interview at any point to add to or delete from the household roster. This option did not require the interviewer to enter a reason for the add or delete, but the interviewer was asked to question the respondent to determine whether the suggested add or delete should be made according to the Census 2000 residence rules.

Input from the respondents’ and interviewers’ perspective

We hired a contractor to conduct cognitive testing on the CEFU instrument for Census 2000 and incorporated many of their suggestions. The contractors characterized the phrase “living or staying” as clearly understood to signify some degree of permanence. Similarly, they reported that respondents’ interpretations of the phrases “most of the time” converged. The respondents’ interpretations of the phrases around the beginning of April” and “as of April 1st” showed more variety, but the respondents generally showed an appreciation of the contingent nature of deciding whom to count as a household member.

They found the respondents did experience some difficulties with the length of the question probes and the structure of the series of items devised to remove people from the household count because they would be counted elsewhere. They also found some sensitivity when respondents with large households were asked to provide census data for members of their household for whom there were no pages on the census form. In addition, the contractor gave us time estimates for the interviews that were in line with our initial projections. Based on these findings, we made some minor changes to the wording and structure of the instrument probes.

Additionally, we conducted agent field trials where we solicited comments from experienced telephone interviewers. They gave us valuable feedback on how to better word and structure the questions, but because our testing ran late, the programmers did not have time to make all the improvements we would have desired.

Results of Coverage Edit Followup for Census 2000

Over 2.4 million households were selected for coverage edit followup, with a little over half of these cases being completed. For each of the completed cases, we looked at the number of times each coverage probe led directly to a name being added or deleted from the household roster.

Number of added persons from completed CEFU cases (Table 1)

Probe topic	Adds
Children	33,246
Relatives	17,088
Non Relatives	10,442
Away temporarily / moving	3,440
No other permanent place	4,436
Interrupt add	83,497
Total adds	152,149

The question about children led to the most name adds, although many more persons were added through the interrupt option. Our analysis showed that the distribution of interruption adds was about one third children, while over half were other relatives, and the remaining about 10% were non-relatives.

Almost one out of every ten CEFU cases contained at least one added person, over 150,000 persons in all. These persons would not have been counted in Census 200 had they not be added through this operation.

Number of deleted persons from completed CEFU cases (Table 2)

Probe topic	Deletes
College student living away?	19,103
Military living away?	2,022
Special places?	5,165
Multiple residences?	16,255
Interrupt deletes	164,368
Interrupt ‘duplicates’	50,654
Total deletes and duplicates	257,567

The questions about college students and persons with multiple residences yielded the most success in making deletions to correct the roster of CEFU cases, although as with the adds, most of the changes were made through the interrupt option. We have no data to

explain the reasoning for persons deleted through the interrupt option.

More than out of every ten CEFU cases contained at least one deleted person, over 250,000 persons in all. These roster corrections insured that we would not count these persons at households where they were not residing on Census day.

Conclusions

The coverage edit followup operation for Census 2000 was successful correcting the roster for one out of each five completed cases. Over 400,000 persons were added or deleted from the initial household roster as reported on their original mailback Census 2000 form. Many of these corrections were made following the asking of coverage probes questions, although the majority were not.

The question about children led to the most name adds, while the questions about college students and persons with multiple residences yielded the most success in making deletions to correct the roster of CEFU cases.

More research is needed in order to plan the coverage edit followup operation for the 2010 Census. Work on the internal evaluation for coverage edits will be completed this year. We will be looking at many things, including how the edit failure rates and case completion rates varied across various subgroups.

We will also be evaluating the effectiveness of the coverage probes used during Census 2000. Question ordering, wording, length, as well as the number of questions will all be evaluated. The goal of this work is to find ways to increase the completion rate because we can only improve coverage and eliminate confusion about household size if we are able to complete these coverage edit followup interviews.

Acknowledgments

I would like to thank James Treat and Leslie Anne Brownrigg of the Census Bureau for their helpful suggestions.