

Within Household Coverage Results of the 2004 Census Test Form

Kyra Linse, David Sheppard
U.S. Census Bureau
Decennial Statistical Studies Division

Abstract

The Census Bureau is currently conducting tests to evaluate the methods and questions to be used for the 2010 Census. One line of research is related to within household coverage. This is our attempt to ensure that the correct persons are included on each census form. In other words, everyone should be counted once, only once, and in the right place according to the official Census residence rules.

The mailback form is the form sent to most known housing units in the United States. This form contains the questions necessary to collect information for all the persons living in that housing unit. This form includes residence rules and two coverage questions to identify possible problems in within household coverage.

If the household does not return the mailback form, the household is then contacted by an enumerator to conduct a Nonresponse Follow Up (NRFU) interview. After the responses from mail return and NRFU are collected, a Coverage Research Follow Up (CRFU) interview is conducted on cases that either had inconsistent information or where the roster is likely to be incorrect. This paper will look into the history of the within household coverage questions, how they are used in selection of the coverage followup universe, and the results of the coverage questions in the 2004 Census Test.¹

Key Words: Coverage Improvement, Within Household Omissions, Erroneous Enumerations

1. History of Coverage Questions/ Coverage Follow Up

A coverage followup operation has been used in the last two decennial censuses primarily to improve within household coverage. The identification of cases for followup has changed between 1990 and 2000, and will most likely change for the 2010 Census.

¹ This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The view expressed on statistical and methodological issues are those of the authors and not necessarily those of the U.S. Census Bureau.

The two primary sections on the census form that are related to within household coverage are the residence rules instructions and the coverage questions. The responses to coverage questions sometimes led to a followup in 1990; in 2000, the coverage questions directed the enumerator to make the appropriate corrections on the form before submitting it.

On the 1990 census questionnaire, at the end of the mailback questionnaire, there were two household level coverage questions. The undercount coverage question asked about people not listed because the respondents were unsure about including them. The overcount coverage question asked if the respondent listed anyone they were unsure about including such as visitors staying there temporarily or people who usually live somewhere else.

▪ 1990 Undercount Coverage Question:
Did you leave anyone out of your list of persons for Question 1a on page 1 because you were not sure if the person should be listed – for example, someone temporarily away on a business trip or vacation, a newborn baby still in the hospital, or a person who stays here once in a while and has no other home?

- Yes, please print the name(s) and reason(s).
 No

▪ 1990 Overcount Coverage Question:
Did you include anyone in your list of person for Question 1a on page 1 even though you were not sure that the person should be listed – for example, a visitor who is staying here temporarily or a person who usually lives somewhere else?

- Yes, please print the name(s) and reason(s).
 No

If the respondent had problems deciding who should be included on the mailback questionnaire and marked one or both of these two coverage questions, the household was included in a coverage edit followup interview. The coverage followup enumerators resolved the situation by either making changes to the roster (by adding people to and/or deleting people from the form) or by deciding the form was correct as originally completed.

In the 2000 Census, there were coverage questions on the enumerator questionnaires for Nonresponse

Followup (NRFU), but there were no coverage questions on the mailback questionnaire.

For NRFU in 2000, the coverage questions were asked after all 100 percent data items were collected. Both questions were at a household level. The undercount coverage question asked if the enumerator missed any children, including foster children, anyone away on business or vacation, any roomers or housemates, or anyone else who had no other home. If there were people identified, they were added to the household by the NRFU enumerator.

- Census 2000 Undercount Coverage Question:
I need to make sure I have counted everyone who lived or stayed here on April 1, 2000. Did I miss—
 - any children including foster children?
 - anyone away on business or vacation?
 - any roomers or housemates?
 - anyone else who had no other home?
 - Yes
 - No

The overcount coverage question asked if any of the people listed were away at college, in the Armed Forces, in a nursing home, or in a correctional facility. If there were any people identified, the cancellation box was marked (a box connected to a single person to delete that person) and the people were considered removed from the household. Therefore, there was no need to follow-up households that marked either of the two coverage questions because it was assumed the enumerator resolved any questions and people were added or deleted correctly.

- Census 2000 Overcount Coverage Question:
The Census Bureau has already counted certain people, so I don't want to count them again here. On April 1, 2000, were any of the people you told me about—
 - away at college?
 - away in the armed forces?
 - in a nursing home?
 - in a correctional facility?
 - Yes
 - No

As a result of the undercount and overcount coverage questions, we added 77,050 people and deleted 83,160 people. Among the people recorded as adds, 46.6 percent were non-Whites, 57.9 percent were young people (ages 0 to 24), 56.5 percent were males, and 51.2 percent were renters. These groups are

traditionally undercounted, however, we cannot infer that these two coverage questions are good for improving the differential undercount because of the proper data was not collected for a conclusive evaluation.

2. 2004 Census Test

The 2004 Census Test was conducted in two sites: Queens, New York and three counties in Georgia. There were approximately 175,000 housing units in the mailout/mailback universe and 25,000 housing units in update/leave area.

The goals of our research in the 2004 Census Test were to improve the way we communicate the residence rules to respondents and to help us determine the usefulness of our coverage questions for the 2010 Census. We wanted to develop the coverage questions so that we can accomplish two things. First, we wanted to determine if we could identify a better way to enumerate certain types of people who historically have coverage problems, like college students, without doing followup. In addition, we needed to determine if following up with households based on their responses to coverage questions is an effective way of improving coverage and decreasing the differential undercount.

The first step toward improving coverage was updating the coverage questions. For each question there were goals based on previous versions of the question and the data it collected. Initial drafts of the 2004 questions were developed through a series of steps including: brainstorming, review of the results of 2000 coverage question evaluations, review of the 2000 census results in terms of undercount and overcount, and other Census processes. The questions were then cognitively tested, resulting in the final questions used in the 2004 Census Test.

2.1. Undercount Coverage Question

The undercount coverage question aimed at identifying potential people not included in the count (undercount). It was asked after the size of the household has been determined on the mailback forms. For NRFU, it was asked after the roster was created by enumerators. It is asked for the entire household.

The undercount coverage question asked if there was anyone not included in the count or not listed who sometimes lives or stays here. The type of people we wanted them to consider included those whom the respondent does not consider to be a household member, were unsure about whether to list, and/or people deliberately left off because of reasons such as

fear of deportation, eviction, or loss of benefits. The final question used was:

2. Are there other people who live or stay at this place part of the time but are not permanent residents, such as live-in employees or children in joint custody?

No

Yes

Number of people at this place part-time =

A response was considered a ‘Yes’ if the yes checkbox was checked or a number other than zero was written in the number of people at this place part-time box.

We are hoping to develop the best question to identify housing units with undercoverage and omissions that need to be followed up without causing confusion or inaccurate data collection.

2.2 Overcount Coverage Question

The overcount coverage question was asked for each person. This was a change from the previous years since it is now done for person 1 through person 6 instead of at the household level. We ask the overcount coverage question to identify people who potentially were misreported. For example, the residence rules instructions state that a person like a college student who lives away should not be included, but the college student was included in the household anyway; the overcount coverage question should be marked to identify this error. The final question wording used was:

10. Does this person sometimes live or stay somewhere else?

No Yes—Mark all reasons that apply.

<input type="checkbox"/> To attend college	<input type="checkbox"/> To be closer to work
<input type="checkbox"/> To stay at a seasonal or second residence	<input type="checkbox"/> For a child custody arrangement
	<input type="checkbox"/> For any reason not listed

A response was considered a ‘Yes’ if the yes checkbox or any of the reason checkboxes were checked.

Based on those persons with a Yes response to the overcount coverage question in combination with the data from followup, we hope to determine what living situations may need to be added to the residence rules instructions or if there are certain characteristics of people or households we should followup more often to remove the person from this household.

3. Results of Coverage Questions for the 2004 Census Test

3.1 Undercount Coverage Question Results

The results of the undercount coverage question allow us to make a preliminary assessment of the number of possible missed people on the household roster. It may also give us some information on the respondent’s comprehension of the question by comparing the answers received by mode (self response mailback vs. personal visit for nonresponse followup). However, it is the results of the CRFU that will really inform the effectiveness of these questions. See reference Krejsa for results.

Tables 1 and 2 show the results of the Undercount Coverage Question for the mailback universe by which boxes they checked crossed by how they filled the Number of Part-Time Residents count box. The results have been split by site.

Overall, 6.2% (1500/24,241) of the Georgia mailback universe had possible missing people on the household roster identified based on the undercount coverage question. (Checked the Yes answer box (1,216+24) or entered a number in the possible part-time residents count (172+88)).

Overall, 4.9% (3,095/62,626) of the New York mailback universe had possible missing people on the household roster identified based on the undercount coverage question. (Checked the Yes answer box (2,235+63) or entered a number in the possible part-time residents count (479+318))

Tables 1 and 2 also show where there was possible confusion with the question since the answer we expected was a ‘Yes’ check box and a number together or only the ‘No’ check box. The confusion by respondents is most apparent in columns 2 and 4 since either an answer box was not checked or both boxes were checked.

For Georgia, 4.7% of households (1,146/24,241) did not mark the question in the way it was intended to be answered. They either did not check yes or no (859), checked both ‘yes’ and ‘no’ (24), filled in a number when checking a ‘no’ checkbox (172), or they filled in a ‘yes’ checkbox and did not provide us with a number other than zero (79+12). For New York, that percentage was even higher at 6.9% (4334/62,626).

It seems entering the number of possible part time residents is confusing to the respondents or they are not connecting it to the undercount coverage question.

Of those who checked a yes box in Georgia, 8.3% did not fill in the number correctly (79+12+8+4/1,216+24). Again in New York, the percentage is even higher at 12.6% (223+25+36+6/2,235+63).

Another sign that some respondents are not understanding the question is that they are skipping the question entirely. As seen on the table, 3.0% of the households did not provide an answer in Georgia and 5.0% of households in New York. (Data not shown in a table) NRFU did not collect the possible number of part-time residents, but instead collected the names. It used the exact same question.

For NRFU (data not shown in table), the number of nonresponses dropped at both sites to 1.0% in Georgia and 1.3% in NY, which is significant between modes (self-response vs. interview) at a 10% confidence level. The number of households with possible missing people on the roster also dropped to 2.2% in Georgia and 2.4% in New York which is also significant between modes. The difference in nonresponse and less possible missing rosters members can be attributed to NRFU being conducted by an enumerator who can answer questions and explain the residence rules if necessary. Not having to identify a number of possible part-time residents may have contributed to the drop in nonresponses as well.

3.2 Overcount Coverage Question Results

The goal of the overcount coverage question was to identify possible erroneous enumerations in the household roster by asking if that person lived somewhere else sometimes. For the 2004 test, a goal was to identify the common reasons people stay somewhere else that result in possible erroneous enumerations.

Again, we separated mailback from NRFU results to see if the mode of collection has any influence on the responses. And again, it is the results of the CRFU that will really inform on the effectiveness of these questions. See reference Krejsa for results.

Tables 3 and 4 show the person level results of the overcount coverage question for the mailback forms. Persons 7 and higher are not included since they were not asked this question.

In Georgia, 5.8% of the persons (3,371/58,615) were identified as possible erroneous enumerations on the roster by either having checked the Yes answer box (358+9) or checked a reason box (3,004). These people were in 2,375 of the households, so 9.8%

(2,375/24241) of the households would need followup based on the overcount coverage question.

In New York, 7.1% of the persons (10,813/153,295) were identified as possible erroneous enumerations on the roster by either having checked the Yes answer box (1,071+78) or checked a reason box (9,664). These people were in 7,218 of the households, so 11.5% (7,218/62,626) of the households were identified as needing followup based on the overcount coverage question.

In Tables 3 and 4, many of the boxes show the question or the way to answer may not have been totally understood. It is interesting that 22.0% of those who checked a reason did not check the Yes or No answer boxes in New York (2,124/9,664), but the respondent may not have understood they needed to do both. It is more confusing to determine why 6.8% of the respondents who checked a reason also checked just the No box in New York (660/9,664).

From reviewing the type of reasons being marked, 47.3%(358+9+980/2,882+26) of those who checked the 'yes' box did not provide a reason or checked the reason 'Other' in Georgia and 53.7% (1,071+78+3,166/7,887+142) did the same in New York for mailback. This is something the Census Bureau has to look at to determine if there is another category that we should be listing or using the CRFU data, if these are not true overcoverage persons.

As seen in the tables, 3.2% of the persons in Georgia and 4.8% in New York left the question completely blank. This could be possible confusion about the question or how to answer, but could also be linked to the respondent not knowing the answer for other members of the household.

For NRFU, the overcount coverage question did not have a 'Yes' response choice. The respondent could either answer one of the reasons or answer 'No'. Only 2.4 % of NRFU persons in Georgia and 1.8 % in New York were possible erroneous enumerations based on selecting a reason in the overcount coverage question. For NRFU persons, 95.8% in Georgia would not require a followup versus the 91.0% for mailback cases. In New York, it is the same pattern, 95.7% of NRFU persons versus 88.2% for mailback.

In Georgia, the people who were possible erroneous enumerations were within 792 of the households, meaning only 4.8% of the NRFU household would need followup based on the overcount coverage question. This is lower than mailback.

For New York, it follows the same pattern. The people who were possible erroneous enumeration fell into 2,417 households, meaning only 4.2 % of the NRFU households would be sent to followup based on the overcount coverage question. This is also lower than mailback.

Tables 5 and 6 show the distribution of the reasons selected for both mailback and NRFU by site. The multiple row is for persons who selected more than one reason. Overall, the reason selected most was “Other” with just over 30% of those giving a reason selecting it at both sites. This tells us that there is a reason people live somewhere else part of the time that we have not properly identified yet. This may also indicate that respondents are not understanding the listed reasons. It is important to try and determine why they are selecting ‘Other’ to make our targeting better.

Besides ‘Other’ always being the most popular reason, respondents select more than one reason the least for all sites and modes. Within site and mode the patterns differ.

In Georgia mailback, the second largest percentage of people answered “Child Custody” 23.3% of the time. In NRFU, this was also the second most popular reason selected. Of the single reasons, “Closer to Work” was chosen the least with it being answered only 10.1% in mailback and 13.5% in NRFU. The distribution of the reasons selected are the same between modes for Georgia.

In New York, the second largest percentage of people answered “seasonal or second home” in mailback at 30.4%. In NRFU, this was the most popular reason selected even over the reason “Other”. Of the single reasons, “Child Custody” was chosen the least of the time with it being answered only 4.6% in mailback and 3.6% in NRFU. In New York, the distribution of reasons selected differ between modes.

Multiple responses were marked by 2.7 % of those with a response in Georgia and 3.7% of those with a response in New York. This was more prevalent in mailback than NRFU. In the two sites for mailback, a pattern developed that out of the 20 different multiple combinations answered, seasonal or second home was checked with something else for 58.4% of the people who checked more than one reason. This could be an indication that people away part time consider their other residence a second home even if it is for work or college.

These numbers become more relevant based on the results of CRFU to see if selecting a certain reason more likely indicates a true erroneous enumeration or not. This will help decide which of the persons that check ‘Yes’ to the overcount coverage question should be in followup in the future.

References and Acknowledgements

Childers, D.; Krejsa, E.; and Sheppard D. (May 2003), “Study Plan for Evaluation of the 2004 Census Test: Evaluation of Residence Rules and Coverage Question, Evaluation 6.” 2010 Census Test Memoranda Series, Chapter: 2004 Census Test, No. 3.

Krasko, N; Darsey, A.; James J.; and Ingold, J. (August 2003), “2004 NRFU Instrument Specifications with Updates 4.2”

Sheppard, D. (February 2003), “Coverage Edit Followup Operation” Census 2000 Evaluation I.1.

Krejsa, E. (October 2005) “Results of Coverage Followup in the 2004 Census Test” Joint Statistical Meeting Contributed Paper.

Nguyen, N and Zelenak, M. (June 2003) “Coverage Gains from Coverage Questions C-1 and C-2 on Enumerator Completed Questionnaires for Census 2000” 2000 Evaluation I.5

Table 1: Results of Undercoverage Question for Mailback forms only - GEORGIA

# Of Part-time Residents	Marked ‘Yes’ Box Only		‘Yes’ and ‘No’ box marked		Marked ‘No’ Box Only		Neither box marked		Total	
	N	%	N	%	N	%	N	%	N	%
Blank	79	.3	8	< .1	20,895	86.2	717	3.0	21,699	89.5
Zero	12	.1	4	< .1	1,075	4.4	54	.2	1,145	4.7
(1+)	1,125	4.6	12	.1	172	.7	88	.4	1,397	5.8
Total	1,216	5.0	24	.1	22,142	91.3	859	3.5	24,241	100

Table 2: Results of Undercoverage Question for Mailback forms only – NEW YORK

ASA Section on Survey Research Methods

# Of Part-time Residents	Marked 'Yes' Box Only		'Yes' and 'No' box marked		Marked 'No' Box Only		Neither box marked		Total	
	N	%	N	%	N	%	N	%	N	%
Blank	223	.4	36	< .1	54,418	86.9	3,098	5.0	57,775	92.3
Zero	25	<.1	6	< .1	1,887	3.0	128	.2	2,046	3.2
(1+)	1,987	3.2	21	< .1	479	.8	318	.5	2,805	4.5
Total	2,235	3.6	63	.1	56,784	90.7	3,544	5.7	62,626	100

Table 3: Person level results of overcoverage – mailback - GEORGIA

Main Check Boxes	No Reason Marked		Reason Marked		Total	
	N	%	N	%	N	%
'Yes' Only	358	0.6	2,464	4.2	2,822	4.8
'No' Only	53,365	91.0	98	0.2	53,463	91.2
'Yes' and 'No' marked	9	< 0.1	17	< 0.1	26	< 0.1
Neither box marked	1,879	3.2	425	0.7	2,304	3.9
Total	55,611	94.9	3,004	5.1	58,615	100

Table 4: Person level results of overcoverage – mailback – NEW YORK

Main Check Boxes	No Reason Marked		Reason Marked		Total	
	N	%	N	%	N	%
'Yes' Only	1,071	0.7	6,816	4.4	7,887	5.1
'No' Only	135,154	88.2	660	0.4	135,814	88.6
'Yes' and 'No' marked	78	< 0.1	64	< 0.1	142	0.1
Neither box marked	7,328	4.8	2124	1.4	9,452	6.2
Total	143,631	93.7	9,664	6.3	153,295	100

Table 5. Distribution of Reasons for living somewhere else - GEORGIA*

Reasons	Mailback		NRFU		Total	
	N	%	N	%	N	%
Attend College	306	10.2	175	15.9	481	11.7
Seasonal Home	615	20.5	250	22.8	865	21.1
Closer to Work	303	10.1	148	13.5	451	11.0
Child Custody	702	23.3	254	23.2	956	23.3
Other	980	32.6	257	23.4	1,237	30.2
Multiple	98	3.3	13	1.2	111	2.7
Total	3,004	100	1,097	100	4,101	100

* Percents are run by mode (i.e. the % under College Mailback is the percent of those who marked a reason in mailback)

Table 6. Distribution of Reasons for living somewhere else – NEW YORK*

Reasons	Mailback		NRFU		Total	
	N	%	N	%	N	%
Attend College	1,120	11.6	512	15.4	1,632	12.6
Seasonal Home	2,935	30.4	993	30.0	3,928	30.3
Closer to Work	1,550	16.0	737	22.2	2,287	17.6
Child Custody	440	4.6	119	3.6	559	4.3
Other	3,166	32.7	928	28.0	4,094	31.5
Multiple	453	4.7	27	0.8	480	3.7
Total	9,664	100	3,316	100	12,980	100

* Percents are run by mode (i.e. the % under College Mailback is the percent of those who marked a reason in mailback)