

1994 NATIONAL CENSUS TEST: QUANTITATIVE COMPARISONS OF COVERAGE

Joan M. Pausche

United States Bureau of the Census, Washington, DC 20233

Key Words: Residence Rules, Reinterview, Nonsampling Error

The 1994 National Census Test was conceived in an effort to determine the best means for improving the respondent's ability to respond correctly.

The primary objective of the project is to test various questionnaire design modifications, such as changes pertaining to the coverage and roster questions, by measuring the extent to which they affect coverage. The study seeks to examine gross coverage error, which includes erroneous enumerations and missed persons. The secondary objective is to determine the extent to which design changes affect response.

The Census Test was focused on inadvertent respondent roster errors which stem from a lack of understanding of the residence rules. The study neither attempted to address deliberate concealment of persons nor intended to measure effects of large-scale content changes.

Two test questionnaires were mailed to a national sample of housing units. The questionnaires contained differences in the wording and sequence of the roster and/or its probes, the coverage edit questions, and the usual-home-elsewhere (UHE) question. Descriptions of the questionnaires are as follows:

Improved 1990 Questionnaire

This form used the 1990 rostering method with minor content and format modifications to the roster question (Step 1 on form). The form was an 8-page respondent-friendly booklet. The whole-household-usual-home-elsewhere (WHUHE) question (Step 6 on form) was moved from the roster page to the back of the questionnaire. The coverage edit questions (Step 4 and Step 5 on form) were reworded and placed after the demographic questions yet before the housing questions. Step 4 identifies persons who might inadvertently be left off the form by asking, "Is there anyone who should be counted here that you haven't listed, for example, someone temporarily away on

a business trip or someone who stays here once in a while and has no other home?" Step 5 identifies persons already listed who might have another residence by asking, "Does anyone listed on this form have another place they usually live, for example, someone visiting or staying temporarily?"

The WHUHE question identifies households for which every member has a usual home elsewhere by asking, "Is this your household's usual address, or is it a temporary address (such as a vacation home)?" It also asks for the WHUHE address.

Extended Roster

The form was a 12-page respondent-friendly booklet. This form's roster question greatly expanding the boundaries of who should be included. The roster is created by the initial question (Question 1 on form) and additional roster probes (Questions 2 through 5 on form).

In order to screen persons included because of the less restrictive rostering method, a set of four individual UHE indicator questions was integrated into the population questions for each person on the roster (Questions 6-9 for Person 1 on form). This enabled the Bureau to identify who should not be counted at the residence rather than allowing the respondent to make this determination.

The first UHE question asks if there is another place where the person lives or stays, the second asks for a description of the place, the third asks where the person lives and sleeps most of the time, and the last asks for the UHE address.

A prenotice letter, a reminder postcard, and a replacement questionnaire were mailed out across all test panels to improve response. In addition, the questionnaire envelopes had a printed message informing the respondents that participation in the survey is required by law. This message had been associated with higher response rates in a previous census test, the 1992 Appeals and Long-form Experiment (ALFE).

After mail-return questionnaires had been checked-in, a telephone reinterview was conducted to identify respondent roster errors, such as erroneous exclusions and inclusions of persons. The reinterview questionnaire contained numerous and detailed probes that enabled the respondent to provide a more accurate list of household members relative to the concise roster questions that appeared on the test questionnaires. The residence status of each person as determined by reinterview was compared to the residence status of the individual based on the mail-return questionnaire.

The reinterview was designed to be conducted with the original respondent, whenever possible. Of completed interviews, approximately 96 percent were actually conducted with the original respondent.

SAMPLE DESIGN AND LIMITATIONS

A stratified sample of clusters of two housing units was selected and the housing units within each cluster were randomly assigned to the two panels. Two strata were formed by delineating census tracts based on the following demographic characteristics, which are associated with census coverage error: race, Hispanic origin, and tenure.

The Low Coverage Area (LCA) stratum had a high proportion of minority persons and renters; the High Coverage Area (HCA) contained the residual. Within the LCA stratum, about 42 percent of the persons were Black, 27 percent were Hispanic, and 57 percent were renters. Within the HCA stratum, 5 percent of the persons were Black, 4 percent were Hispanic, and 27 percent were renters.

The two strata were formed using the FASTCLUS procedure in SAS to mathematically delineate the census tracts based on minority rate and percent renters per tract. The procedure assigned each tract to one of two strata such that the variability of minority percentage and renter percentage per tract were minimized in each of the final two strata.

The mailout sample for the LCA stratum was 19,000 questionnaires per panel, or a total of 38,000 questionnaires. The mailout sample for the HCA stratum was 3,000 questionnaires per panel, or a total of 6,000 questionnaires. The test was designed to measure a 0.5 percentage point minimum difference in gross error rate in the LCA stratum and a 0.7 percentage point difference in gross error rate in the HCA stratum. The mailout sample assumed that

approximately 50 percent of the deliverable forms in the LCA stratum and 65 percent in the HCA stratum would be returned based on the total number of questionnaires mailed (actual response rates were 51 and 75 percent, respectively).

About 51 percent of the mail-returned questionnaires were subject to reinterview. Completed interviews totalled 9,574 out of 11,402 attempts.

The test was limited to respondent-administered, mail-returned questionnaires. Thus results and conclusions do not apply to respondents enumerated by personal interviews during followup operations. In addition, approximately 25 percent of the housing units were erroneously excluded from the sampling frame during the automated sample selection process. These excluded units were distributed throughout the frame although a slightly larger proportion were located in the HCA stratum. The restricted frame included at least some portion of every census tract in the study universe and the pattern of excluded units was not correlated with any demographic variables.

In addition, housing units in which every person had a UHE, as determined by the questionnaire data, were excluded from the analysis. These cases were excluded since the focus of this study was on within-household rostering errors rather than errors in the identification of households in which everyone usually lived somewhere else.

RESULTS

All estimates are weighted to reflect the sample design. Variance estimates were produced using the jackknife estimation method for a stratified cluster sample.

Gross Error Rate and Components

The primary measure of within-household coverage for this study is gross error rate, which is defined as the sum of the within-household missed persons (omissions) and the within-household erroneously enumerated persons, divided by the number of true residents, as determined by reinterview.

This estimate can be viewed as measure of response error, which includes both response variance and bias. Response variance is the variability among repeated measures and response bias is the difference between the observed value and the true value of the item of interest. Note that although the gross error has

components of both response variance and bias, it might not include the total bias, in terms of the mail-return residence status compared to the reinterview residence status.

A person was classified as an omission if they were identified as a true resident of the housing unit (i.e., those who should have been enumerated at the housing unit) based on reinterview, but were either not listed on the census questionnaire or classified as a non-resident based on responses to the UHE questions on the Extended Roster Form. Persons who listed themselves as respondents at the end of the questionnaire but were not listed anywhere else on the form, and were classified as residents during reinterview were considered omissions.

A person was classified as an erroneous enumeration if they were identified as a non-resident based on reinterview but were identified as residents based on roster data on the Improved 1990 Form or UHE indicator data on the Extended Roster Form.

In some cases, mail return data were not sufficient for determining the residence status of persons identified on the mail-return form. Since their correct residence status could have been determined through a followup interview, these persons were classified as resident or non-resident of the sample housing unit based on the reinterview in order to simulate the outcome of a census. No coverage errors were charged to persons with undiscernible mail-return residence status, that is, these persons were **not** classified as erroneous enumerations or omissions.

For the Improved 1990 Form, all persons listed in the coverage probes had undiscernible mail-return residence status. For the Extended Roster Form, persons for whom the UHE indicator data was incomplete or missing were classified as undiscernible. This applies to cases in which more than 8 persons were listed on the roster. Since the form was designed to contain demographic and UHE data for up to 8 persons, additional persons on rosters with more than 8 did not have corresponding person boxes in which to provide UHE data.

Table 1 contains gross error rate and component estimates at the national and stratum levels by form type. Error rates are shown as percents and standard errors are shown in parentheses.

TABLE 1: Gross Error Rates and Components

	EE Rate	Om Rate	GE Rate
National			
Impr 90	1.48 (0.21)	0.58 (0.19)	2.06 (0.26)
Ext Rost	0.92 (0.13)	0.79 (0.14)	1.71 (0.18)
HCA			
Impr 90	1.27 (0.30)	0.52 (0.16)	1.79 (0.34)
Ext Rost	0.59 (0.16)	0.66 (0.17)	1.24 (0.23)
LCA			
Impr 90	2.07 (0.22)	0.79 (0.12)	2.86 (0.25)
Ext Rost	1.90 (0.21)	1.20 (0.15)	3.10 (0.25)

Based on the analysis, there was no significant difference between gross error rates between panels at the national or stratum levels. However, the percentage of persons with undiscernible residence status based on the mail-return form and not included in the error estimates was considerably larger for the Extended Roster Form compared to the Improved 1990 Form. Although the vast majority of the persons with undiscernible status on the Extended Roster Form were classified as residents during reinterview, the amount of error introduced by UHE information for these persons, had it been provided by the respondent, is unknown.

The erroneous enumeration rate for the Extended Roster Form at the national level, as shown above in Table 1, is significantly smaller than the erroneous enumeration rate for the Improved 1990 Form (SE of difference=0.269).

Yet, the omission rate for the Extended Roster Form in the LCA stratum is not significantly different from the omission rate for the Improved 1990 Form (SE of difference=0.177).

Effectiveness of Extended Roster Form UHE Indicators

The UHE indicators were included on the person data pages of the Extended Roster Form so that the census would be able to determine the UHE status for each person listed on the mail-return roster without relying on followup for the information. The performance of the UHE indicators is examined in detail since, from an operational viewpoint, an excessive followup workload could be considered a "fatal flaw" in the form, regardless of the form's performance as measured by the gross error rate or completion rate.

Based on study data, an estimated 13.53 percent (SE=2.55 percent) of the persons on the form classified as non-residents of the housing unit based on the UHE indicators were classified as residents, and thus omissions, based on the reinterview.

Conversely, approximately 1 percent (SE=0.14 percent) of the persons classified as residents of the housing unit based on UHE information were classified as non-residents, and thus erroneous enumerations, based on the reinterview.

Of the persons on the Extended Roster Form, 10.71 percent (SE=0.59 percent) had incomplete or missing UHE information and thus had undiscernible questionnaire residence status. During a census, edit followup would be conducted to determine these persons residence status. Based on the results of additional analysis, there was no apparent clustering effect of persons requiring followup with respect to questionnaires; an estimated 17.6 percent of the Extended Roster Forms would require followup. This compares to an estimated 2.3 percent of the Improved 1990 Forms that would require followup of persons listed in the coverage questions.

Effectiveness of Extended Roster Form Probes

The purpose of the roster probes is to encourage the respondent to provide data for persons who have an attachment to the household. Although some non-residents may be brought in by these probes, one would expect that the majority of persons included as a result of the probes would be residents of the sample housing unit. One would also expect that the residence status of the persons brought in by the probes would be correctly classified by the UHE indicators so that non-residents could be removed from the roster.

To determine the proportion of persons who were included by each of the roster probes but for whom the UHE indicators failed to correctly determine the residence status, the analysis explored the enumeration status of persons listed on the form cross-classified by where their name first appeared on the form.

Although Probe 3 brought in more persons than the other probes (2.91 percent of the persons on the Extended Roster Form), an estimated 37.69 percent (SE=3.12 percent) were correctly classified as non-residents. In addition, 17.74 percent (SE=3.87 percent) of the persons listed as a result of Probe 3 were classified in error by the UHE indicators and 38.64 percent (SE=3.16 percent) could not be classified.

An estimated 41.85 percent (SE=7.57 percent) of the persons included by Probe 4 were correctly classified as residents but over 43 percent (SE=7.75 percent) could not be classified. The majority of persons included by Probe 5 were non-residents. The high residence status error rates and high percentage of persons with unknown residence status is common to all of the roster probes.

Effectiveness of Improved 1990 Form Coverage Questions

The placement and wording of the coverage questions on the Improved 1990 Form was different from the coverage questions on the 1990 Census form. These changes in the coverage questions were expected to help create a more accurate roster by identifying residents who were initially missed and non-residents who should have been removed from the roster.

Contrary to the expected results of the first coverage question, 57.07 percent (SE=6.28 percent) of the persons listed in Step 4, which asks for anyone else who should have been listed, were non-residents and only 34.69 percent (SE=5.87 percent) were residents.

The second coverage question was similarly ineffective in that 66.70 percent (SE=6.39 percent) of the persons listed in Step 5, which asks for anyone already listed who has a UHE, were residents and only 29.94 percent (SE=6.11 percent) were non-residents. These results are contrary to the anticipated coverage benefit of these questions.

In addition, only 0.63 percent (SE=0.08 percent) of persons listed on the form were included in Step 4. Another 0.30 percent (SE=0.08 percent) were listed

incorrectly in Step 4 since they had already been listed on the roster. Only 0.76 percent of the persons listed on the form were identified in Step 5 as having a UHE, with an additional 0.19 incorrectly included in Step 5 since they had not appeared anywhere else on the form.

Completion Rates

The mail-return completion rate is defined as the number of questionnaires checked-in divided by the number of questionnaires mailed-out, excluding those classified as Post Master Returns (PMRs). For this study, a housing unit was classified as a PMR if both questionnaires, the initial and the replacement questionnaires, were returned as undeliverable.

Note that the 1994 Census Test completion rates are not directly comparable to census mail response or mail return rates. The response rate is the percentage of forms checked in based on all forms mailed out and the return rate is the percentage of forms checked in based on occupied housing units in the mail-out. In comparison, the completion rate should be less than or equal to the return rate since the completion rate denominator could contain vacant units that were not classified as PMRs.

Table 2 shows weighted completion rate estimates and their standard errors at the national and stratum levels by form type. Completion rates are shown as percents and standard errors for the estimates are shown in parentheses.

TABLE 2: Completion Rates

	Weighted Completion Rate Estimate %
National Total	74.29 (0.46)
Impr. '90	74.83 (0.62)
Ext. Roster	73.76 (0.63)
HCA Stratum	78.89 (0.56)
Impr. '90	79.41 (0.76)
Ext. Roster	78.37 (0.77)
LCA Stratum	54.39 (0.28)
Impr. '90	55.05 (0.37)
Ext. Roster	53.73 (0.37)

Concurrent with sample design assumptions, the completion rate for the LCA stratum was significantly lower than the completion rate for the HCA stratum; 54.39 percent compared to 78.89 percent, respectively, as shown in Table 2.

The estimated completion rate differences between panels was significant only in the LCA stratum. Yet it is important to note that differences in the length of the questionnaires (the Extended Roster Form was 4 pages longer than the Improved 1990 Form) may have affected the completion rates.

CONCLUSIONS

In general, one of the objectives of the 1994 National Census Test is to determine the best means for improving the respondent's ability to understand what roster information we seek and thus, respond correctly. To attain this, the questionnaire should not contain unclear, confusing, or problematic items. The overall goal is to identify a rostering method that not only increases within-household coverage and reduces the amount of enumeration error, but also maximizes questionnaire response.

Regarding within-household coverage, there was no significant difference between the two panels in gross error rates.

Both the coverage questions on the Improved 1990 Form and the roster probes on the Extended Roster Form were problematic. High residence status misclassification rates and large percentages of persons with unknown residence status were common to all of the roster probes on the Extended Roster Form. Overall, the percentage of persons with incomplete or missing UHE data was 10.71 percent.

With regard to the coverage questions on the Improved 1990 Form, the majority of the persons listed in Step 4 (which asks for anyone else who should have been listed) were non-residents.

The majority of persons listed in Step 5 (which asks for anyone already listed who has a UHE) were residents. Both of these results are contrary to the anticipated coverage benefit of these questions. In addition, the coverage gain from these questions was minimal.

Lastly, the estimated completion rate differences between panels was significant only in the LCA stratum. However, differences in the length of the questionnaires (the Extended Roster Form was 4 pages

longer than the Improved 1990 Form) may have affected the completion rates.

In general, both forms had small gross error rates which indicates that both forms are effective in producing an accurate roster. However, the coverage questions on the Improved 1990 Form as well as the roster probes and UHE indicator questions on the Extended Roster Form were problematic and need further revision. The results of this test provided a valuable basis for designing the 1995 Census Test mail-return form.

ACKNOWLEDGEMENTS

The author thanks Deborah Griffin and Nicholas Alberti of the Census Bureau for their helpful suggestions.

REFERENCES

Cochran, William G. Sampling Techniques. John Wiley and Sons. New York. 1977.

Fay, R. "Draft Documentation of VPLX." Unpublished document. U.S. Bureau of the Census. January 17, 1990.

Jones, C. "Enumeration and Residence Rules for the 1990 Census." 1990 Decennial Census Policy Memorandum No. 12. October 15, 1987.

This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau.