

2010 Decennial Census Program – Coverage Measurement Program

Spring 2006 Advisory Committee Meetings

April 5, 2006

**Donna Kostanich/David Whitford
U.S. Census Bureau**

This document is being provided to the Census Bureau's Advisory Committees prior to upcoming meetings. It is preliminary in nature and in the early stages of development. As such, it is subject to revision. Our intent in making this working document available at this time is to inform ongoing discussions related to 2010 Census planning and the Coverage Measurement Program.

QUESTIONS

1. Our CCM plans emphasize data collection and matching in order to improve determination of census day residence and to identify and resolve any duplicate census enumerations. Do you think that there are any coverage measurement issues that are of greater importance than this?
2. Our plans for CCM in 2010 may be ambitious given the expanded scope. Any feedback you could provide on how we should prioritize this work to get the best results?
3. Any other thoughts on which coverage estimates are the most important?

Introduction

A coverage measurement program is essential for measuring the effectiveness of a census in terms of assessing coverage, determining whether the strategic goals of the census are met, and identifying ways to improve future census designs and operations.

The overall coverage goal for the 2010 Census is to “improve the accuracy of the census data, especially the coverage of the population and housing inventory, for all geographic levels and demographic subgroups,” (Angueira 2004). Two targeted strategic goals aimed towards achieving the overall coverage goal are to: 1) reduce duplication of persons and housing units, and 2) reduce net coverage error (Waite 2004).

The coverage measurement program for past censuses has gone by different names: in 1980 it was called the Post-Enumeration Program (PEP), in 1990 it was called the Post-Enumeration Survey (PES), and in 2000 it was called the Accuracy and Coverage Evaluation Survey (A.C.E.). A major objective of all these previous coverage measurement efforts was to provide results that could potentially be used to improve the census counts. To this end the PEP, PES, and the A.C.E. programs were designed primarily to provide estimates of the net coverage error in their respective censuses. A main purpose for census coverage measurement in 2010 (2010 CCM) is to understand how coverage error relates to census operations, in order to understand how those operations might be improved to reduce coverage error in subsequent censuses. This change in purpose implies a need to provide estimates of the components of coverage error (omissions and erroneous inclusions), rather than just the net coverage error.

The goals for the 2010 CCM are more ambitious than they have been in the past, even though there are no plans to use the results for improving the census counts. Thus, major changes in the coverage measurement methodology and operations are being developed and tested in activities leading up to the 2010 Census, namely the 2006 Census Test and the 2008 Dress Rehearsal.

Plans for the 2010 Census Coverage Measurement

The broad goals of the 2010 CCM are to (Singh 2003):

- begin producing measures of coverage error components (omissions and erroneous enumerations)
- produce measures of coverage error components for demographic groups and for key census operations
- continue to provide measures of net coverage error

Obtaining estimates of the components of coverage error--separate estimates of census omissions and erroneous enumerations--is our highest priority. Obtaining estimates of net error will continue to be necessary since they are needed to estimate omissions. In addition to producing estimates of coverage error for persons in housing units, we also plan to produce estimates of

coverage error for housing units and for persons in housing units by whether the housing unit was missed or enumerated.

Our original plans for 2010 included producing coverage error estimates for persons in group quarters and possibly for group quarters facilities. However, as we began developing the details of the design and methodology needed to meet our goals for persons and housing units, it became clear that we lacked the lead time and resources needed to design, develop, and implement the group quarters component of the coverage measurement program. Consequently, the 2010 CCM program will not provide coverage estimates for persons in group quarters or for group quarters facilities.

A more detailed discussion of each of the broad goals for 2010 follows.

Measures of Coverage Error Components

We plan to produce estimates of coverage error components (omissions and erroneous enumerations), along with estimates of net coverage error. Duplicates in the census are a special type of erroneous enumeration. We also plan to produce person coverage estimates by housing unit coverage status. The results should help answer questions such as the following (Kostanich, and others, 2004):

- How many persons and housing units were missed?
- How many persons were missed because their housing unit was missed?
- How many persons were missed in housing units that were correctly enumerated?
- How many housing units were enumerated in the wrong location?
- How many persons were enumerated at the wrong residence?
- How many persons were enumerated in the wrong location because their housing unit was geocoded to the wrong location?
- How many persons and housing units were enumerated more than once; that is, duplicated?
- How many housing units were enumerated erroneously for other reasons, such as newly constructed units completed after Census Day?
- How many fictitious persons were enumerated?
- How many persons were erroneously enumerated for other reasons: born after April 1, died before April 1, or moved into the sample address after April 1?
- How many housing units were enumerated with incorrect occupied or vacant status?

Explore Relationships Between Components of Coverage Error and Census Operations

As noted above, we would like to understand the relationship between measured components of census coverage error and particular census operations. For example, we would like to determine how housing unit duplication rates differ by type of census enumeration area or by the original source of the unit's address. Our original plans included determining correct or

erroneous enumeration status for any record that ever had a chance of being counted in the census, regardless of whether it was eventually included in the census. Due to resource limitations, these plans have been scaled back somewhat. As a result our ability to understand the relationship between census omissions and the census operation that may have incorrectly excluded these records from the final census file will be limited. We will be able to examine erroneous census inclusions introduced by specific census operations such as nonresponse followup or coverage followup.

Coverage Errors for Demographic Subgroups

Producing estimates of net coverage error will continue to be an important coverage measurement goal. Our plans for 2010 include providing measures of the net coverage error and the differential net coverage error with respect to demographic characteristics. We are also interested in understanding the components of coverage error for these same demographic groups.

Coverage Errors at the Subnational Level

Estimates of coverage error (net and components) at the subnational level are important to the extent that they may provide information about how particular census operations concentrated in those areas might be improved. In the past, synthetic estimation has been used to produce estimates of net coverage error for small geographic areas. Synthetic estimates are required when there is not sufficient sample available for direct estimation.

Estimating net coverage error for small subnational areas (county-level or lower) is not as important in 2010 as it was in previous censuses, because we will not be producing results for possible adjustment of the census counts. However, our attempt to relate coverage error, including components, to census operations could lead to forming subnational estimates using models or something different than the traditional synthetic estimation methodology. There are no plans to produce estimates of the number of people erroneously enumerated in a specific state or county. We may also produce national-level estimates of the number of people erroneously enumerated in a state or county.

Survey Design Issues

In developing our survey design plans for 2010, two important issues have been identified: contamination and respondent burden. Both of these issues involve the relationship between coverage measurement operations and census operations.

A fundamental assumption underlying the coverage measurement estimation methodology is independence between the coverage measurement survey and the census. Violating this assumption can lead to serious deficiencies in the coverage measurement results. Additionally, census results themselves could be affected if coverage measurement operations are allowed to

occur while census operations are taking place. When coverage measurement affects the census, contamination is said to exist. The potential for such contamination exists with the coverage measurement person interview (PI) and the census coverage followup operation (CFU) since their scheduled dates in the field overlap. Contamination of the census results could occur in cases where the PI occurred before the CFU interview. However, since the coverage measurement sample is very small compared to the size of the census, the effect on census results would be minimal. There could, however, be a large bias in the coverage estimates, because 1) the independence assumptions are violated, or 2) census coverage could be different between the coverage measurement sample areas and those areas not in the CCM sample.

In other cases, the coverage measurement results could be affected if the CFU interview occurs before the PI. This latter situation could also lead to bias in the coverage estimates. A strategy to reduce the potential bias would be to use information on whether or not a housing unit was in the CFU (and possibly the reason it was in the CFU) as a post-stratification variable or as an independent variable in logistic regression models for producing the estimates.

Another potential problem between the PI and the CFU is respondent burden. Both interviews are attempting to collect similar information about which persons were living at a particular address on Census Day. Even though there are important differences between the interviews (for example, the CFU interviewer is given access to the census data, while the PI interviewer starts from scratch), those differences may not be evident to the respondent and may result in confusion or the respondent's refusal to complete the second interview.

Since the same operations were involved in both issues, various alternatives addressing both issues simultaneously have been proposed and evaluated based on expected availability of staff resources, timing issues, and perceived operational difficulties. Alternatives considered to address the respondent burden issue were: conducting a combined CFU/PI interview for CFU households and using the PI interview as a substitute for the CFU interview (or the CFU interview as a substitute for the PI interview). Alternatives to address the contamination issue (and the possible adverse affect on coverage measurement estimates) were: delaying the PI until after CFU operations are completed, not doing the census followup interview in CCM sample blocks, and using census data which would not include people added to the census after a certain point in time (prior to the start of PI interviewing) to produce the coverage measurement estimates. This latter option has also been referred to as "truncating the census." See Davis and Bell (2005) for more details on the respondent burden issue and see Bell (2005a, 2005b) for more details on the contamination issue.

A decision has been made for 2008 and 2010 to begin the coverage measurement PI after the CFU has been completed. Additionally, the PI interview will be allowed to occur for all sample cases even those that may have had a CFU interview. For 2006, the two interviews will be allowed to occur during the same time period. This will provide an opportunity to study the potential contamination of census results, possible effects on the coverage measurement data, and to see if the response rates for either operation are affected. Not delaying the PI in 2006 will

afford us the opportunity to finish coverage measurement activities earlier so that we will be able to incorporate improvements into those activities for the 2008 Dress Rehearsal.

Details of the 2010 Census Coverage Measurement

Listed below are some details on the design and operations of the 2010 census coverage measurement program:

- We will use capture/recapture sampling and estimation methods, including the dual system estimator.
- Our independent sample will consist of clusters of census blocks, comprising about 300,000 housing units throughout the U.S. and about 15,000 housing units in Puerto Rico. The 2010 sample will be allocated similar to the way the sample was allocated for the 2000 A.C. E.
- Data collection will focus on people who are living or staying at the sample address on Interview Day, including those people who have moved into the address since Census Day. Information about certain people who have moved away from the sample address since Census Day will also be collected, when appropriate.
- A followup interview will be conducted for persons for whom we don't have enough information to determine their Census Day residence.
- The initial person interview and the person followup interview will use different approaches to gathering information on when moves took place. The person interview will use a more qualitative approach and the person followup interview will use a quantitative approach, gathering more exact information on the timing of any moves that may have occurred.
- The initial person interview will be automated, but the person followup interview will not be automated. In both interviews we will attempt to gather enough information to allow us to determine where each person should have been counted in the census and where they may have been counted in the census. That is, we will collect information on alternate addresses.
- There will be an automated search for person matches between the coverage measurement survey sample and the census and an automated search for duplicate census enumerations across the country, with limited followup to determine the actual duplication status and Census Day residence. The country-wide search for duplicates will compare census enumerations in the CCM sample areas to all census enumerations.
- Matching operations within the sample block cluster and a surrounding ring of blocks will consist of both automated and clerical searches, along with appropriate field followup.
- We will attempt to match census enumerations with deficient or missing person name and minimal data characteristics to persons in the coverage measurement survey sample.
- We will identify person matches and duplicate census enumerations at alternate addresses collected in the initial or followup interview by searching the block containing the alternate address, along with surrounding blocks.

- Coverage estimates for persons in housing units, housing units, and persons in housing units by housing unit enumeration status will be produced. No coverage estimates will be produced for persons in group quarters or for group quarters facilities.

Coverage Measurement Goals of the 2006 Census Test and the 2008 Dress Rehearsal

Based on the problems identified with the 2000 A.C. E. and the new goals set for the 2010 Census coverage measurement program, four broad areas of research were identified for the 2010 census coverage measurement program (Kostanich and others, 2004):

- improving determination of Census Day residence
- improving techniques to detect and resolve duplicate census enumerations
- developing methods for measuring the components of coverage error
- improving techniques to estimate net coverage error

Coverage measurement activities for the 2006 Census Test will focus on data collection and matching operations needed to address the first three of these research issues. The 2006 Census Test will include our first attempt to collect the additional information we need to determine where a person should have been counted on Census Day (previous coverage measurement efforts determined only whether a person was enumerated correctly in the coverage measurement survey sample areas). An evaluation being conducted as part of the 2006 Census Test will focus on our ability, as a result of all of the changes being made to the coverage measurement survey design, questionnaires, and computer and clerical matching operations to determine a person's residence on Census Day.

In an attempt to detect and resolve census duplicate enumerations in the 2006 test, we are testing enhancements to the matching operations, including an automated search for person matches between the coverage measurement survey sample and the census, as well as a search for duplicate census enumerations within each site. There will be followup to determine actual duplicate status and the actual Census Day residence of people identified as possible duplicates.

We will also attempt to match census enumerations having a missing or deficient person name or minimal data characteristics to persons in the coverage measurement survey sample. This enhancement is necessary in order for us to be able to more accurately estimate the coverage error components.

Although we do not plan to evaluate the coverage of the 2006 Census Test, we plan to use the data collected in 2006 to determine whether or not we are collecting adequate and consistent data needed to produce component and net coverage error estimates. In the 2008 Dress Rehearsal, the focus will shift towards testing the various steps in the estimation process and continuing to search for ways to improve our methodology for estimating net coverage error.

For the 2006 coverage measurement survey sample, there will be no independent listing of housing units (the census listing will be used). However, the CCM roster of persons in these housing units will be independent from the census.

Another evaluation being conducted for the 2006 test will look for evidence of contamination of census results, potential effects on the coverage measurement data, and the potential impact on response rates as a result of conducting the coverage measurement PI interview during the same time period as the CFU interview.

The 2008 Dress Rehearsal will also provide our first opportunity for testing the operations and methodology for other aspects of the coverage measurement program, including: independent housing unit listing, matching, and followup and producing estimates of person coverage by housing unit coverage status. No major difficulties were found with these operations in the 2000 A.C.E.

REFERENCES

- Angueira, Teresa, (2004), "Action Plan: 2010 Research and Development Planning Group on Coverage Improvement," 2010 CENSUS PLANNING MEMORANDA SERIES No.19, U.S. Census Bureau, February 17, 2004.
- Bell, William, (2005a), "Contamination of Coverage Follow-Up Results by the Coverage Measurement Interview – Implications for Estimation of Net Error and Error Components," DSSD 2010 CENSUS COVERAGE MEASUREMENT MEMORANDUM SERIES #2010-E04, April 21, 2005.
- Kostanich, Donna, David Whitford and William Bell, (2004), "Plans for Measuring Coverage of the 2010 U.S. Census," American Statistical Association Joint Statistical Meetings, 2004 Proceedings of the Section on Survey Research Methods.
- Kostanich, Donna, David Whitford (2005), "Discussion of the Proposal to Reduce Respondent Burden for Households in both Census Coverage Followup and Census Coverage Measurement," DSSD 2010 CENSUS COVERAGE MEASUREMENT MEMORANDUM SERIES #2010-A07, November 29, 2005.
- Singh, Rajendra P., (2003), "2010 Census Coverage Measurement – Goals and Objectives (Executive Brief)," DSSD 2010 CENSUS COVERAGE MEASUREMENT MEMORANDUM SERIES #2010-A01, October 1, 2003.
- Waite, Preston Jay, (2004), "Targets for 2010 Census Strategic Goals," 2010 CENSUS

DECISION MEMORANDUM SERIES No. 4 (Revised), May 17, 2004.