

SUPPORTING STATEMENT
U.S. Department of Commerce
U.S. Census Bureau
2010 Census Coverage Measurement Initial Housing Unit Followup
OMB Control # 0607-XXXX

B. Collections of Information Employing Statistical Methods

1. Universe and Respondent Selection

The 2010 Census Coverage Measurement (CCM) sample design has been developed to support the various objectives of the program, which includes estimating erroneous enumerations and omissions in addition to net error for the 2010 Census. The CCM is designed to measure the coverage of housing units and persons, excluding group quarters and persons residing in group quarters. Remote areas of Alaska are out-of-scope for CCM.

The CCM sample consists of two parts. The Population, or P sample, and the Enumeration, or E sample, have traditionally defined the samples for dual system estimation, a statistical technique for measuring net coverage error. The P sample and the E sample measure the same housing unit and household population. However, the P-sample operations are conducted independent of the census. The E sample consists of census enumerations in the same sample areas as the P sample. For net error estimation, after matching with the census lists and reconciliation, the P sample provides information about the population missed in the census, whereas the E sample provides information about erroneous census inclusions. The correct enumeration rate and match rate are used to estimate the true population total via dual system estimation. A new objective of the CCM in 2010 is to measure the components of census errors, that is, the omissions and erroneous census enumerations.

The CCM sample design comprises a number of distinct processes from forming block clusters, creating the sampling frame, selecting sample block clusters, to eventually selecting addresses for the P and E samples. After the CCM block clusters are selected, an address list is created independent of the census within each CCM sample block cluster. The approximate CCM listing workload is 12,500 block clusters for the nation and 600 for Puerto Rico. Overall, approximately one million housing units are listed--950,000 in the nation, and 50,000 in Puerto Rico. Finally, after selecting the CCM sample addresses for interviewing, there are approximately 300,000 housing units for the nation and 15,000 for Puerto Rico. The national sample is distributed among the 50 states and the District of Columbia roughly proportional to population size, although there are slight increases in the sample for small states and for American Indian Reservations.

Table 1 summarizes the National and Puerto Rico universe sizes from Census 2000, along with the CCM expected listing workloads and the P-sample sizes. The E-sample size is expected to be about the same as the P sample.

Table 1: 2010 CCM Universe and Sample Housing Unit Summary

Geography	2000 Census	Expected Listing Sample Size in 2010	Expected P-sample Size in 2010
U.S.	115,904,641	950,000	300,000
Puerto Rico	1,418,476	50,000	15,000
Total	117,323,117	1,000,000	315,000

The CCM sample has three phases of sampling. In the first phase, we form block clusters from contiguous collection blocks. The block clusters in each state are classified by size into mutually exclusive and relatively homogeneous groups known as sampling strata. These strata are based on the block cluster size and whether the block cluster is located on an American Indian Reservation. The four major strata, which are the same as those used for the 2000 Accuracy and Coverage Evaluation (A.C.E.) first phase of sampling, are (1) block clusters with 0 to 2 housing units (small stratum), (2) block clusters with 3 to 79 housing units (medium stratum), (3) block clusters with 80 or more housing units (large stratum), and (4) block clusters on American Indian Reservations with three or more housing units (American Indian Reservation stratum). Using 2000 census data, the medium and large strata are further split into renter and owner block clusters, resulting in up to six sampling strata being formed in each state and Puerto Rico.

Block clusters with 80 or more housing units are selected with higher probability than medium block clusters in this phase because housing units in large block clusters will be subsampled in a later operation. This will bring the overall probability of selection – the inverse of the sampling weight – for housing units in these block clusters more in line with the overall selection probabilities of housing units in medium block clusters. Block clusters in the renter strata are selected at a higher rate than block clusters in the owner strata. Within each sampling stratum, block clusters are sorted and a systematic sample is selected with equal probability.

In the second phase, small block clusters are subsampled to improve operational efficiency and to reduce costs, while also attempting to minimize the variance impact. Conducting interviewing and followup operations in block clusters of this size is more costly per housing unit than in medium or large block clusters. Using housing unit counts from the independent list and the updated census address list, within each state we re-stratify the small block clusters selected in the first phase by size and select systematic samples from each stratum with equal probability. All block clusters from the small sampling stratum with 10 or more housing units based on the updated information are retained. All block clusters from the small sampling stratum which are on American Indian country are also retained. (American Indian country includes American Indian Reservations and associated trust lands, as well as the American Indian statistical areas.)

In the third phase of CCM sampling, we select a subsample of housing units within large block clusters. If a block cluster contains 79 or fewer housing units, all the housing units are included in the CCM sample. For block clusters with 80 or more housing units, a subsample of these housing units is selected to facilitate data collection in the field and to reduce the impact of intracluster correlation on the variance. This phase of sampling results in more similar overall selection probabilities for housing units because the large block clusters will have a higher

probability of selection during the first phase. This subsampling will be done by forming groups of adjacent housing units, called segments. A systematic sample of segments within each block cluster will be selected. All housing units in the selected segments will be included in the CCM sample.

For the third phase of CCM sampling, the sampling frame for the P-sample housing units is the result of the CCM initial housing unit matching and followup operation. The intent of this housing unit operation is to resolve differences between the independent housing unit list and an early census housing unit list. This can result in housing units being removed from the independent listing, but no units can be added to the independent listing. In addition to sending the P sample to the Person Interview, a sample of census units which were missed during the independent listing operation will be sent to the Person Interview. While not part of the P sample, these census units are likely to be in the E sample. The P-sample persons result from the person interviewing in the P-sample housing units.

The sampling frame for the E-sample housing units consists of the housing units in CCM sample areas from the Census Unedited File (CUF), which is available after the P sample is selected. While these two samples are selected at different points in time, we attempt to geographically overlap them to the extent possible. The E-sample persons are the census enumerations in the E-sample housing units that have at least two characteristics, of which name can be one; these are referred to as census- data defined enumerations.

2. Procedures for Collecting Information

The CCM Initial Housing Unit Followup Form, Form D-1303, is used by interviewers to resolve discrepancies from matching the CCM Independent Listing and the census address lists. This form is translated into Spanish for use in Puerto Rico. The Spanish form is Form D-1303 (PR). Interviewers will contact a member (or proxy, as a last resort) of each housing unit being followed up to ensure the accuracy of the address lists. Interviewers will be provided the Housing Unit Reference list (see Attachments G and H) to aid in following up housing units. The interviewer will give each respondent in the 50 states and the District of Columbia a copy of the Form D-31(CCM-IHUFU), U.S. Confidentiality Notice; in Puerto Rico, the interviewer will give each respondent a copy of the Form D-31(CCM-IHUFU)PR, Puerto Rico Confidentiality Notice. The Confidentiality Notice explains that any information given to the Census Bureau will be held in strict confidence. It also informs each respondent that participation is mandatory. For the Initial Housing Unit Followup Quality Check (QC), approximately 10 percent of all housing units followed up will be identified for a Quality Check. The QC interviewers will locate the housing units identified for the quality check and compare the information collected by the initial interviewer to what they see on the ground. A single attempt will be made to contact a respondent for the quality check. If no one is available, the quality check will be completed by observation. Block clusters not passing the QC will be 100 percent verified to ensure the data quality of the Initial Housing Unit Followup. QC interviewers will use the Initial Housing Unit Followup Form, Form D-1380, for block clusters in the United States, and the D-1380(PR) for block clusters in Puerto Rico.

3. Methods to Maximize Response

The Initial Housing Unit Followup contains the minimum number of questions necessary to obtain the data required for the 2010 CCM. The interviewer will make up to three attempts to obtain an interview. The interviewer will explain the reason the Census Bureau is conducting this operation, and respondents will be informed of their legal responsibility to answer the questions. In addition, respondents will be assured that their answers are confidential.

4. Testing of Procedures or Methods

The Census Bureau developed the CCM approach for measuring the coverage of the population in the decennial census. It was used in the 2000 Decennial Census, and the approach was updated and refined for the 2008 Census Dress Rehearsal. However, the Initial Housing Unit Followup was dropped from the Dress Rehearsal due to budget cuts. Instead, a 2008 Initial Housing Unit Followup Operational Test was conducted in the Dress Rehearsal sites, San Joaquin County, California, and South Central North Carolina. CCM Regional Managers from Charlotte and Seattle conducted interviewer training for Census headquarters employees unfamiliar with the IHUFU operation. These staff were paired with someone familiar with the IHUFU questionnaire and goals of the operation, who helped identify what went well and identified areas of concern with the form, training, or materials. The clusters selected for the IHUFU operational test were clusters that would have gone for follow-up had the operation occurred. These targeted housing units contained the changes from the 2000 questionnaire.

5. Contacts for Statistical Aspects and Data Collection

Gia F Donnalley
Coverage Measurement Data Collection Operations Branch
Decennial Statistical Studies Division
U.S. Census Bureau
301-763-4370

Definition of Terms

Components of Coverage Error – The two components of census coverage error are census omissions (missed persons or housing units) and erroneous inclusions (persons or housing units enumerated in the census that should not have been). Examples of erroneous inclusions are persons or housing units enumerated in the census that should not have been enumerated at all, and persons or housing units enumerated more than once (duplicates).

Net Coverage Error – A measure of the difference between census omissions and erroneous inclusions. A positive net error indicates an undercount, while a negative net error indicates an overcount.

For more information about the Census 2010 Coverage Measurement Program, please visit the following page of the Census Bureau’s website:

<http://www.census.gov/cac/www/pdf/coverage-measurement-program.pdf>

List of Attachments

- A. U.S. Initial Housing Unit Followup Form, Form D-1303
- B. Puerto Rico Initial Housing Unit Followup Form, Form D-1303 (PR)
- C. U.S. Initial Housing Unit Followup Quality Control Form, Form D-1380
- D. Puerto Rico Initial Housing Unit Followup Quality Control Form, Form D-1380 (PR)
- E. U.S. Confidentiality Notice, Form D-31(CCM-IHUFU)
- F. Puerto Rico Confidentiality Notice, Form D-31(CCM-IHUFU)PR
- G. U.S. Housing Unit Reference List, Form D-1303.REF
- H. Puerto Rico Housing Unit Reference List, Form D-1303.REF(PR)
- I. Definitions for Initial Housing Unit Followup U.S. Flashcard, Form D-1028.4(CCM)
- J. Definitions for Initial Housing Unit Followup Puerto Rico Flashcard, Form D-1028.4(CCM)PR