SUPPORTING STATEMENT
U.S. Department of Commerce
U.S. Census Bureau
2010 Census Coverage Measurement Independent Listing
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 produce a general purpose sample to support the various objectives of the program, which includes the new objective of 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 Sample, P sample, and the Enumeration Sample, E sample, have traditionally defined the samples for dual system estimation. Both 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 using dual system estimation which is used to estimate net error. Using a less restrictive definition, the E sample is used to measure the erroneous enumerations for the new objective of component errors. Omissions are estimated by using the less restrictive definition of erroneous enumerations and net error.

The CCM is a multi-phase sample designed to measure the net coverage and components of coverage for the household population and housing units in the 2010 Census. 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 for 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, the 2010 coverage measurement sample is 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 size from Census 2000, along with the CCM expected listing workloads and the P-sample size. The E-sample size is expected to be the same as the P sample.

Table 1: 2010 CCM Universe and Sample Housing Unit Summary

Geography	2000 Census	Expected Listing Sample Size	Expected P-sample Size
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 of the CCM sampling, 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, bringing 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 from 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 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, we re-stratify the small block clusters selected in the first phase within each state 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 intraclass 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 at 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 which 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 which have at least two characteristics, which name can be one; these are referred to as census-defined enumerations.

2. Procedures for Collecting Information

The CCM Independent Listing Form, Form D-1302, known as the Independent Listing Book (ILB), is used by Listers to canvass every street, road, or other place where people might live in their assigned block clusters in the 50 states (excluding remote Alaska) and the District of Columbia to construct a list of housing units. The ILB is translated to Spanish for use in Puerto Rico. The Spanish form is Form D-1302 (PR). Listers will contact a member (or proxy, as a last resort) of each housing unit to ensure all units at a given address are identified, identify the type of housing unit (single-family, multiunit, mobile home, or trailer), identify the number of apartments occupied or vacant in multiunits, and for a mobile home parks, the number of mobile homes, trailers, and empty trailer lots/sites in the park. They also identify the location of each housing unit by assigning map spots on block cluster maps provided with their assignment materials. Listers will provide each respondent with the confidentiality notice, Form D-31 (CCM-IL) in the 50 states (excluding remote Alaska) and District of Columbia, or Form D-31 (CCM-IL)PR in Puerto Rico. The response rate for the Independent Listing Operation is

expected to be 100%, given that a completion by observation is allowed as a last resort.

For the Independent Listing Dependent Quality Check (DQC), a random starting point from each completed block cluster will be identified to start the selection of 12 housing units from the block cluster for verification. DQC listers will locate the housing unit identified as the starting unit on the ground, and then compare the next 12 housing units they see on the ground to what is listed in the ILB. If there are fewer than 12 housing units in the block cluster, the entire block cluster will be verified. Block clusters not passing the DQC will be 100 percent verified to ensure the data quality of the Independent Listing.

At the completion of the Independent Listing, the ILBs will be keyed to construct the CCM housing unit list needed for subsequent CCM operations.

3. Methods to Maximize Response

The ILBs contain the minimum number of questions necessary to obtain the data required for the 2010 CCM, and 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.

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 – Reflects 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. Independent Listing Book, Form D-1302
B. Puerto Rico Independent Listing Book, Form D-1302 (PR)
C. U.S. Confidentiality Notice, Form D-31(CCM-IL)
D. Puerto Rico Confidentiality Notice, Form D-31(CCM-IL)PR