



**UNITED STATES DEPARTMENT OF COMMERCE**  
**Economics and Statistics Administration**  
**U.S. Census Bureau**  
Washington, DC 20233-0001

July 9, 2008

DSSD 2010 CENSUS COVERAGE MEASUREMENT MEMORANDUM SERIES #A-19

MEMORANDUM FOR Frank A. Vitrano  
Chief, Decennial Management Division

From: David C. Whitford (*signed*)  
Chief, Decennial Statistical Studies Division

Subject: Overview of 2010 Census Coverage Measurement Program

Attached is a high level overview of the activities planned for the 2010 Census Coverage Measurement (CCM) program. Even though there was not a complete Dress Rehearsal of all CCM activities in 2008, CCM still intends to meet the new goals and objectives as initially given in Singh (2003) and updated in Singh (2005).

Any questions regarding the CCM program should be directed to Donna Kostanich or Magdalena Ramos.

Attachment

cc:

DSSD CCM 2010 Contacts List  
DMD 2010 Planning Coordination

U S C E N S U S B U R E A U

## Overview of the 2010 Census Coverage Measurement Program

July 9, 2008

### I. Background

Through Census 2000, the focus of coverage measurement has been determining net coverage error. The major goal for the 2010 Census Coverage Measurement (CCM) program is to study coverage error in the 2010 Census in order to improve future censuses, meaning 2020 and beyond. This change implies a need to provide estimates of the components of coverage error (omissions and erroneous inclusions), rather than just the net coverage error. Obtaining estimates of net error will continue to be necessary since they are needed to estimate omissions. We will produce these estimates of coverage error for both housing units and persons in housing units. The group quarters universe is out of scope for CCM. Specifics concerning the objectives for the 2010 CCM can be found in Singh (2003) and Singh (2005). It is important to note that coverage measurement will not affect the 2010 Census results.

Coverage measurement was not part of the 2004 or 2005 Census Tests. Testing for CCM began in the 2006 Census Test and continues with the 2008 Dress Rehearsal, however the amount of testing is limited. In the 2006 Census Test, coverage measurement addressed the problems identified with the Census 2000 Accuracy and Coverage Evaluation (A.C.E.) program and began testing new methods to measure the components of coverage error. In 2000, the A.C.E. person interview and person followup interview were ineffective in determining a person's Census Day residence. This resulted in the A.C.E. not being able to adequately identify erroneous enumerations, many of which were found to be duplicates in the census (see Kostanich et al., 2004). Therefore, the sole purpose of coverage measurement in 2006 was to develop and test the CCM survey person phase operations – data collection and matching – with an aim at improving coverage measurement methods. The 2006 CCM plans included conducting an evaluation on whether the new methods were successful in determining a person's Census Day residence. No testing of the CCM housing unit phase operations was conducted. The coverage measurement survey for the 2006 Census Test was not designed to evaluate the coverage of the 2006 Census Test.

Originally the plan was to do a complete dress rehearsal of all CCM person and housing unit operations in 2008 (Vitrano, 2007a). Due to budget shortfalls while the Census Bureau was operating under a continuing resolution at the start of the 2008 fiscal year, it became necessary to cancel many census and CCM operations for the Dress Rehearsal and delay Census Day to May 1, 2008.<sup>1</sup> Additionally, the Census Bureau has decided to descope CCM from the Field Data Collection Automation (FDCA) contract to reduce risk to the 2010 Census operations. The Census Bureau's Technologies Management Office (TMO) has been tasked with the responsibility for the CCM field data collection systems and software development, as well as with developing the CCM Person Interview and Reinterview automated data collection instruments (see Angueira 2008).

---

<sup>1</sup>See Vitrano (2007b) for more on the reduced scope of the 2008 Dress Rehearsal.

## II. Brief Description of 2010 CCM Survey Operations

This section provides a brief description of the major CCM operations. A diagram is included as an attachment which shows the relationship and timing of the CCM operations. The numbered boxes correspond to the descriptions in this section. The ovals on the left hand side of the attachment show key census activities and how they relate to CCM. The diagram also shows those coverage measurement operations that were either not conducted as part of the 2008 Dress Rehearsal or were not fully tested using dress rehearsal data. Although not specifically mentioned, all CCM operations include some type of quality control or validation.

More comprehensive documentation on the 2010 CCM sample design, the survey design, and the estimation process will be available as work progresses.

### 1, 2, 4, 10, 12            CCM Sampling

The 2010 coverage measurement survey sample will be a probability sample of approximately 300,000 housing units in the U.S. (excluding remote Alaska) and approximately 15,000 housing units in Puerto Rico. Two samples are selected to measure census coverage of housing units and the household population: the population sample (P sample) and the enumeration sample (E sample). The P sample is a sample of housing units and persons obtained independently from the census for a sample of block clusters. The E sample is a sample of census housing units and enumerations in the same block clusters as the P sample. The CCM sampling 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.

The CCM primary sampling unit is a block cluster, which consists of one or more geographically contiguous census blocks. Block clusters are formed to balance statistical and operational efficiencies. A stratified sample of block clusters is selected for each state or state equivalent. An independent address list is created for each CCM sample block cluster. This is expected to result in listing approximately one million addresses for the U.S. and Puerto Rico. A double-sampling design is used for small block clusters. This entails selecting a larger sample of small block clusters for which an independent address list is created, and then later selecting a subsample of these block clusters to remain in sample based on additional information. After a confirmed set of potential housing units is determined through the Initial Housing Unit Matching and Followup, the P-sample housing units are identified. In block clusters with a large number of housing units, a subsample of contiguous housing units will be selected for the P sample. This achieves manageable field workloads for CCM Person Interview (PI) and Person Followup (PFU) without having a big impact on reliability.

The source of the E-sample housing units is the Census Unedited File. In block clusters with a large number of census units, a subsample of contiguous census housing units will be selected, attempting to geographically overlap the P and E samples.

### **3. CCM Independent Listing**

Within each selected CCM block cluster, we will conduct a canvassing operation to construct an independent list of all housing units contained therein. Enumerators will canvass every street, road, or other place where people might live, in their assigned block clusters, and construct a list of housing units. They also identify the location of each housing unit by assigning map spots on individual block maps provided with their assignment materials. Since this operation uses a paper-based Independent Listing Book, it does not use GPS technology. If an enumerator is uncertain whether a particular living quarters is a housing unit, they will be instructed to list it. Following the completion of each block cluster, the listing books are keyed for matching against the census Universe Control and Management (UC&M) files (formerly the Decennial Master Address File--DMAF) for the same areas. Completed listing books are subject to quality control, and those that fail are assigned to be reworked.

### **5. CCM Initial Housing Unit Computer Matching**

During CCM Initial Housing Unit Computer Matching, the keyed file of housing unit addresses developed during CCM Independent Listing is computer matched against the UC&M file of living quarters addresses within each sample block cluster and one ring of surrounding blocks. Addresses are assigned one of three possible outcome codes during computer matching: matched, possibly matched, and not matched<sup>2</sup>.

### **6. CCM Initial Housing Unit Before Followup Clerical Matching**

During CCM Initial Housing Unit Before Followup Clerical Matching, the National Processing Center (NPC) matching staff uses computer-assisted clerical matching techniques, along with CCM and census maps, to review and attempt to match, possibly match, or assign not matched codes to addresses presented to them from the Initial Housing Unit Computer Matching. In addition, the NPC matching staff also searches for duplicate addresses. The clerical matching uses a computer-assisted software called the Initial Housing Unit Matching, Review and Coding System (IHU MaRCS). Cases that remain unresolved following this operation are eligible for Initial Housing Unit Followup. Some examples of unresolved cases include: nonmatched CCM or census addresses, addresses identified as possibly matched or possibly duplicated, matched addresses with unresolved housing unit status and addresses that match to group quarters or that match to surrounding blocks.

### **7. CCM Initial Housing Unit Followup**

During CCM Initial Housing Unit Followup, interviewers collect additional information for addresses unresolved after the Initial Housing Unit Computer and Before Followup Clerical

---

<sup>2</sup>Codes for nonmatched units are actually assigned during a pre-processing operation which occurs before Initial Housing Unit Before Followup Clerical Matching.

Matching operations. The CCM Initial Housing Unit Followup operation attempts to collect additional information that might allow a resolution of match codes for any differences between the Independent Listing results and the census UC&M file and also to resolve potential duplicates. The Initial Housing Followup data collection forms will be printed via Docuprint technology. The questions included for each followup case will vary depending upon the reason the case is being sent to followup. This field operation includes a quality control mechanism.

## **8. CCM Initial Housing Unit After Followup Clerical Matching**

The NPC matching staff will use the results of the Initial Housing Unit Followup (using the completed paper questionnaires) to attempt to match unresolved addresses. The result of this operation is files containing match codes for CCM and census units in the sample block clusters. The followup questionnaires will be sent to keying after this operation.

## **11. CCM Person Interview**

For each sample block cluster, we will conduct a CCM Person Interview (PI) for selected housing units. During CCM PI, interviewers use an automated instrument on a laptop to obtain information about the current residents of the sample housing unit, persons living at the housing unit at the time of interview who may or may not have been there on Census Day (nonmovers and in-movers), and certain persons who moved out of the sample housing unit between Census Day and the time of the CCM interview (out-movers). The information collected for each person includes name, sex, age, date of birth, race, relationship, and Hispanic origin. The interviewer also collects information about alternate addresses to establish where people lived on Census Day, according to census residence rules. The PI will also be conducted for some cases not in the P sample in order to obtain information which may help to resolve their census enumeration status earlier than if they were included in the E-sample cases sent to PFU.

The processing of alternate addresses and in-movers' Census Day addresses requires geocoding functionality prior to matching that we did not have in the Census 2000 coverage measurement program. Since addresses are collected during both the PI and the Person Followup interview, geocoding will be required for addresses obtained by each operation at two different times. In-mover and alternate addresses identified during the PI will be geocoded to census collection blocks using automated software. An attempt to clerically geocode those PI respondent-provided addresses that can not be computer geocoded will be made. In-mover and alternate addresses provided by the Person Followup Interview will be clerically geocoded, but not computer geocoded.

This field operation includes a reinterview for quality control purposes.

## **13. CCM Person Computer Matching**

During CCM Person Computer Matching, the person data collected during CCM PI is computer matched against all census enumerations. The person data from the CCM PI is also computer

matched against itself within a sample block cluster to identify duplicate persons in the PI. The person computer matching also searches for census duplicates within the U.S. or Puerto Rico. There will be no matching or duplicate searching between the U.S. and Puerto Rico. Matching and duplicate searches are also conducted at respondent-provided alternate addresses. As a result of computer matching, persons are identified as matched, possibly matched, or not matched.

#### **14. CCM Person Before Followup Clerical Matching**

During CCM Person Before Followup Clerical Matching, the NPC matching staff uses computer-assisted clerical matching techniques, along with CCM and census maps, to review and attempt to match, possibly match or assign not matched codes to person records (linked or not) as a result of computer matching. In addition, clerical matchers conduct clerical searches for duplicate persons. The computer-assisted clerical matching allows the matching staff to determine if a person corresponds to a census enumeration with a missing or incomplete name. It also allows assignment or updating of a person's Census Day residence status. This computer-assisted software is called the Person Matching, Review and Coding System (PER MaRCS). Cases that remain unresolved following this operation are sent to the CCM Person Followup. Some examples of unresolved PFU cases include: nonmatched P-sample person records with a proxy response, nonmatched E-sample person records; possibly matched or possibly duplicated records; matching housing units with different roster in CCM and census (conflicting households), P-sample people with unclassified residence status, in-movers with ungeocoded in-mover addresses, possible matches or possible duplicates at nationwide potential long-distance duplicate addresses (followup at both the in-cluster and long-distance addresses).

#### **15. CCM Person Followup**

During CCM Person Followup (PFU), interviewers contact cases identified in the Person Matching operations as requiring additional information in order to resolve Census Day residence status, enumeration status, match status, or person duplication. The universe for potential followup includes the E sample, the P sample, or suspected census duplicates. The CCM PFU operation collects data that is later used in the Person After Followup Clerical Matching to resolve any differences between the coverage measurement and the census enumeration results. The PFU data collection forms will be created via the Docuprint technology. The questions included for each followup case will vary depending upon the reason the case is being sent to followup. This field operation includes a quality control mechanism.

#### **16. CCM Person After Followup Clerical Matching**

The NPC matching staff use information obtained during PFU from the completed paper questionnaires to attempt to resolve match, residence, enumeration, and duplication status for remaining persons in the coverage measurement P and E samples. The followup questionnaires will be sent to keying after this operation.

## **17. CCM Final Housing Unit Computer Processing**

During CCM Final Housing Unit Computer Processing, housing unit information is processed in order to prepare for the Final Housing Unit Clerical Matching. No computer matching is done during this processing; rather a determination is made as to which housing units will go to the Final Housing Unit Clerical Matching operation. These are generally housing units added to the census after the preliminary list (i.e., the UC&M file) was created or CCM units matched to a census unit that was deleted from the preliminary census list used in Initial Housing Unit Matching.

## **18. CCM Final Housing Unit Before Followup Clerical Matching**

During Final Housing Unit Before Followup Clerical Matching the NPC matching staff use computer-assisted clerical matching techniques, along with CCM and census maps, to attempt to match, possibly match or assign not matched codes to addresses sent from the Final Housing Unit Computer Processing. The clerical matching uses a computer-assisted software called the Final Housing Unit Matching, Review and Coding System (FHU MaRCS). Cases that remain unresolved following this operation are eligible for Final Housing Unit Followup. Some examples of unresolved cases include: nonmatched CCM or census addresses, addresses identified as possibly matched or possibly duplicated, matched addresses with unresolved housing unit status and addresses that match to group quarters or that match to surrounding blocks.

## **19. CCM Final Housing Unit Followup**

During CCM Final Housing Unit Followup, interviewers collect additional information for addresses that are unresolved during the Final Housing Unit Before Followup Clerical Matching. The Final Housing Unit Followup operation attempts to collect information needed to resolve any residual differences between the Independent Listing results and the census. The Final Housing Unit Followup data collection forms will be printed via the Docuprint technology. The questions included for each followup case will vary depending upon the reason the case is being set to followup. This field operation includes a quality control mechanism.

## **20. CCM Final Housing Unit After Followup Clerical Matching**

The NPC matching staff will use the results of the Final Housing Unit Followup from the completed paper questionnaires to match remaining nonmatched addresses. The followup questionnaires will be sent to keying after this operation. This is the final operational step before CCM Estimation.

## **21. CCM Estimation**

The CCM Estimation process consists of several operations, which will ultimately lead to the production of estimates of coverage errors for both housing units and persons in housing units. This includes estimates of net coverage error as well as estimates of the component errors—

omissions and erroneous enumerations. As part of this estimation, we will implement operations to account for missing data and reduce the sampling and nonsampling errors in our estimates.

Like the 1990 Post Enumeration Survey and the 2000 A.C.E., the 2010 CCM will be evaluating net coverage error by using dual system estimation to generate the population estimates of housing units and persons in housing units. For the CCM, we will use logistic regression modeling instead of post-stratification. The logistic regression modeling allows us to reduce the correlation bias in our total population estimates without having to include unnecessary high order interactions as when forming post-stratification cells. Not having unnecessary high order interactions allows us to include additional variables in the model that can potentially help us reduce synthetic error for subpopulation estimates.

For both person and housing unit net error estimation, the major estimation operations are:

- imputations for missing characteristics of the P sample,<sup>3</sup>
- imputations for unresolved enumeration status in the E sample and unresolved match status in the P sample,
- imputations for unresolved Interview Day residence status for person estimation and unresolved Census Day housing unit status for housing unit estimation,
- a weight trimming for influential block clusters,
- logistic regression models and generation of the necessary predictions for estimation.

For person net error estimation, we will implement the following additional operations:

- a noninterview adjustment for the P sample,
- an adjustment if we detect measurement error like we saw in the A.C.E.,
- a correlation bias adjustment to reduce effects of violating dual system model assumptions.

For the measurement of component errors (omissions and erroneous enumerations) in 2010, an enumeration is correct if it is included once and only once anywhere in the census housing unit universe. The implementation of dual system estimation uses a strict definition for correct enumeration that includes correct location as defined by the block cluster and one ring of surrounding blocks. Therefore for the component errors, we need additional processing operations to assign correct and erroneous enumeration statuses to the sample cases. We will also determine and generate estimates by the types of erroneous enumerations (duplicate, fictitious, etc.).

For both person and housing unit component error estimation, we will implement these operations:

- imputations for unresolved enumeration statuses to support erroneous overall and erroneous by types estimates.

---

<sup>3</sup> These characteristic imputations will be used for component error as well.

- a two-stage ratio adjustment estimator to reduce the standard errors for the erroneous enumeration estimate.
- estimate omissions by adding the estimates of erroneous enumerations and net error together. For persons, we will also estimate the number of person omissions by whether their housing unit was included in the census or not.

The last estimation operation will be to generate standard errors of the net error and component error estimates. This operation will use replication methods to generate the standard errors and reflect the sampling and missing status imputation variance.

The Census Bureau is currently compiling a detailed description of the types of coverage statistics that will be produced and will then determine when these might be available for release.

### III. References

Angueira, Teresa (2008). "Decision to Descope Census Coverage Measurement (CCM) Operations from the Field Data Collection Automation (FDCA) Contract," 2010 DECENNIAL CENSUS PROGRAM DECISION MEMORANDUM SERIES No. 21, February 7, 2008.

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.

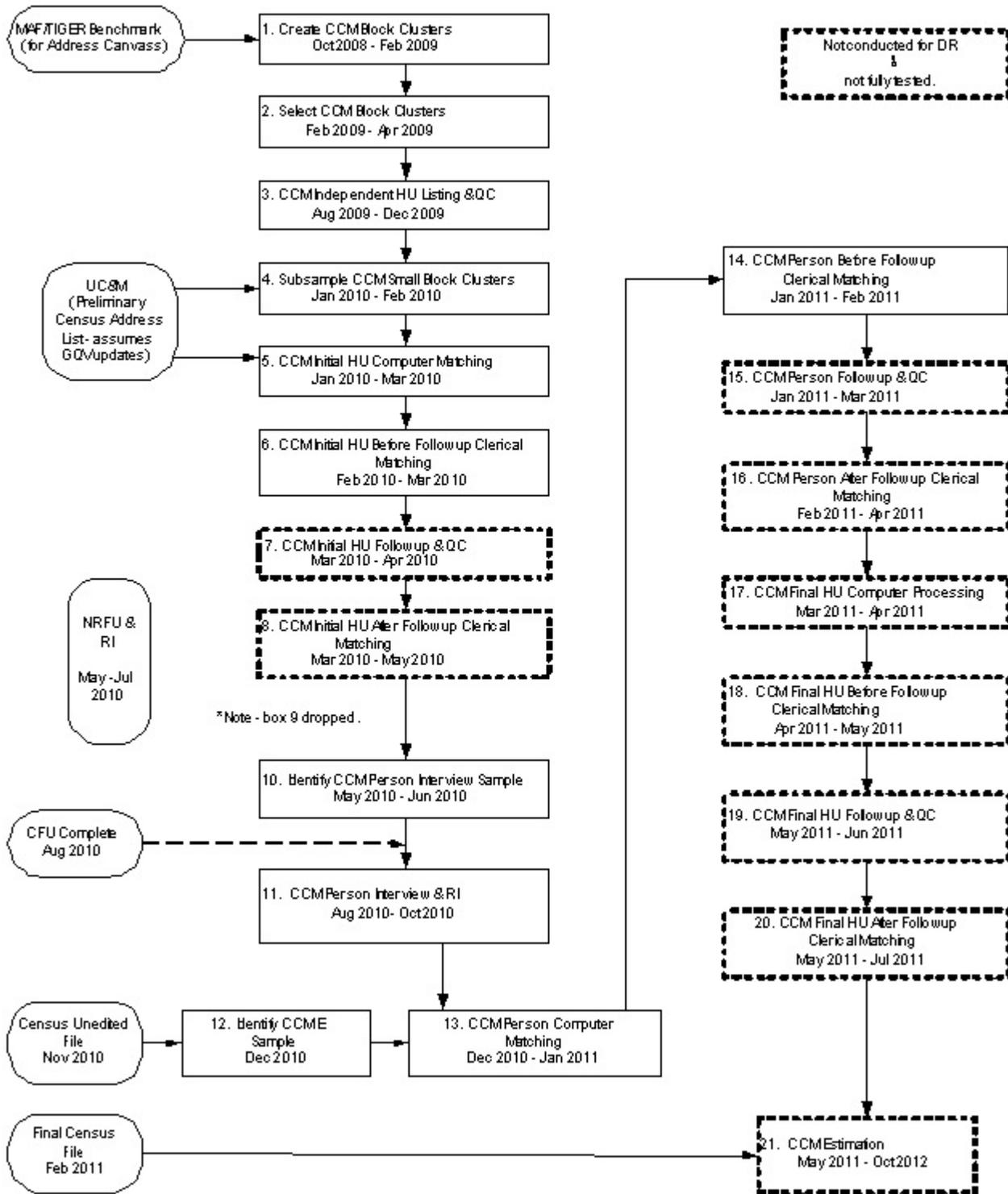
Singh, Rajendra P. (2003). "2010 Census Coverage Measurement – Goals and Objectives (Executive Brief)," DSSD 2010 CENSUS COVERAGE MEASUREMENT MEMORANDUM SERIES #2010-A01, October 1, 2003.

Singh, Rajendra P. (2005). "2010 Census Coverage Measurement – Updated Plans," DSSD 2010 CENSUS COVERAGE MEASUREMENT MEMORANDUM SERIES #2010-A06, November 29, 2005.

Vitrano, Frank (2007a). "2008 Census Dress Rehearsal Detailed Operation and System Plans for Census Coverage Measurement," 2008 CENSUS DRESS REHEARSAL MEMORANDUM SERIES No. 28, June 29, 2007.

Vitrano, Frank, (2007b). "Reduced Scope of the 2008 Census Dress Rehearsal and a One-Month Delay of Census Day," 2008 CENSUS DRESS REHEARSAL MEMORANDUM SERIES No. 50, November 19, 2007.

Diagram: 2010 CCM Major Activities and Milestones



\*The Independent Housing Unit Relisting operation has been dropped because it could not be tested in the dress rehearsal and was only expected to add marginal value. Removing these complications reduces the overall risk to the 2010 CCM program.