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

September 24, 2008

DSSD 2010 CENSUS COVERAGE MEASUREMENT MEMORANDUM SERIES #A-23

MEMORANDUM FOR Documentation

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

Subject: Proposed Census Coverage Measurement Estimates for Net and Component Error

Attached is a paper that was distributed to the Census Advisory Committee of Professional Associations. This served as background information for Thomas Mule's talk at the October 16, 2008 meeting of this committee. As additional background information, the Census Bureau provided the overview memo of the 2010 Census Coverage Measurement Program. The overview memorandum is #A-19 in the DSSD 2010 Census Coverage Measurement Memorandum Series.

Any questions regarding this presentation should be directed to Thomas Mule (301) 763-8322.

Attachment

cc:

DSSD CCM 2010 Contacts List
DMD 2010 Planning Coordination

Proposed Census Coverage Measurement (CCM) Estimates for Net and Component Error

Thomas Mule
Decennial Statistical Studies Division

Background paper for the presentation at the October 16-17, 2008 Meeting
Census Advisory Committee of Professional Associations

Abstract: 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, including 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. This document shows the type of net coverage error and components of coverage error statistics being planned for the 2010 CCM. This includes definitions of the components of coverage error and also includes the currently planned estimation domains.

We are interested in your reaction to:

1. The proposed components of coverage error.

2. The proposed definition of a correct or erroneous enumeration for components of coverage error.

3. The proposed coverage statistics and estimation domains for demographic groupings, geographic areas and census enumeration operations.

This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on issues are those of the author and not necessarily those of the U.S. Census Bureau.

1. Introduction

The purpose of the 2010 Census Coverage Measurement (CCM) program is to evaluate coverage error in the 2010 Census in order to improve future censuses, meaning 2020 and beyond. The CCM is designed to measure the coverage of housing units and persons, excluding group quarters and persons residing in group quarters. The CCM will provide estimates of the net coverage error and the components of coverage error by using a post-enumeration survey. For more information concerning the CCM objectives see Singh (2003) and Singh (2005). Since the CCM is an evaluation, its results will not affect the 2010 Census.

The other principal method to measure coverage is Demographic Analysis. Demographic Analysis represents a macro-level approach, where population estimates are developed by aggregating various types of demographic data. Demographic Analysis can measure national net coverage trends and differences by age, sex, and race (Blacks, Nonblacks). Their program is able to measure the coverage of all persons including the population residing in group quarters. Demographic Analysis will also provide the historical coverage benchmarks to assess change in coverage levels and key differentials (e.g., adult Black men, children). This document focuses on the CCM survey-based estimates being produced. See Judson et al. (2008) for more information on the Demographic Analysis evaluation plans for the 2010 Census.

The 2010 CCM sample design is a large complex survey of 300,000 housing units in the United States (excluding remote Alaska) and 15,000 housing units in Puerto Rico. The CCM survey conducts an independent enumeration of housing units and persons in housing units. The results are matched to census enumerations to identify coverage errors. The CCM consists of five sampling activities, five data collection activities and three matching activities prior to the estimation of coverage error. A high-level overview that shows the relationship and timing of the major CCM activities can be found in Whitford (2008). The CCM program will produce estimates for the United States and Puerto Rico. This background documentation focuses on the coverage estimates being produced for the United States.

Section 2 documents the definitions of coverage errors for the Census Coverage Measurement program. Section 3 describes how the CCM program is implementing these definitions and concepts. Section 4 provides a general approach to the estimation of net census coverage error and the components of census coverage error. Section 5 shows the proposed estimation domains for people in housing units. Section 6 shows the proposed estimation domains for housing units.

2. Definitions for Coverage Measurement Concepts

This section starts by defining the universes that the CCM program will be evaluating. Next, some basic coverage concepts are defined along with the definitions of net coverage error and the different components of coverage error.

2.1 Coverage Universes

Table 1 documents evaluation universes for the CCM. The table lists the two universes for coverage evaluation and the several types of persons and living quarters that are out-of-scope for the CCM. All proposed CCM coverage estimates shown in this document apply only to the two universes that are in-scope.

Table 1: CCM Evaluation Universes

Coverage Universes	Out-of-scope for CCM
1. People in Housing Units 2. Housing Units	1. People in Group Quarter Facilities 2. Group Quarter Facilities 3. People and Housing Units in Remote Alaska 4. People experiencing homelessness or other transient living conditions

2.2 Basic Concepts

The following are definitions of concepts that apply to both coverage of housing units and of persons in housing units. These concepts all refer to enumeration within the housing unit coverage universe, so that everywhere the word "enumerated" is used we could more explicitly say "enumerated in the housing unit coverage universe." Thus, the definition of omission below implicitly includes as omissions persons who were residents of housing units within the coverage universe, but who were enumerated in the census outside the housing unit coverage universe, e.g., in group quarters or in remote Alaska. The definition of erroneous enumeration implicitly includes persons who were residents of group quarters but were enumerated in the census in housing units, etc.

Net Error:	True population – census count.
Omission:	A person or housing unit that should have been enumerated in the census but was not.
Correct enumeration:	A person or housing unit that should have been enumerated in the census and was.
Wrong location:	A correct enumeration that the census has located in the wrong geographic area. Various specific definitions will be used for "correct location" versus "wrong location" (e.g., correct versus wrong state, correct versus wrong county, etc.). When "correct location" is defined as "anywhere in the U.S.," there are no wrong location enumerations.
Erroneous enumeration:	A person or housing unit that should not have been enumerated in the census but was enumerated. This

includes person or housing unit enumeration records that do not correspond to real persons or housing units, such as fictitious persons or demolished housing units.

Duplicate: A type of erroneous enumeration resulting when a person or housing unit that should have been enumerated was included in the census more than once. In such cases one of the enumerations is considered a correct enumeration and the other(s) are considered erroneous due to duplication.

The following define specific other types of erroneous person enumerations:

Discrepant: An enumeration that does not correspond to a real person (e.g., fictitious names, pets, etc.)

Group quarters resident: Enumeration of a person in a housing unit that corresponds to a real person who actually resided in a group quarters.

Born after Census Day: Enumeration of a person who was born after April 1, 2010.

Died before Census Day: Enumeration of a person who died before April 1, 2010.

Other erroneous enumerations: Other types of erroneous enumerations, such as enumerations of (a) U.S. citizens who were working, studying, or living abroad on April 1, 2010, or (b) foreign visitors temporarily in the United States.

Census Imputations: Housing unit or person records in the census requiring imputation. Table 2 describes five different types of imputations in the census. Depending on the amount of information collected, the degree of imputation could range from imputing whether a housing unit even existed to imputing all of the demographic characteristics for part of the household. These characteristics include a person's race, Hispanic-Origin, age (or date of birth), sex, and relationship.

Table 2: Census Imputations by Category

Count Imputation 1. Status Imputation - No information about the housing unit; Imputed as occupied, vacant or non-existent. Those imputed as non-existent are removed from the census files. 2. Occupancy Imputation - Existence of housing unit confirmed but no information as to occupancy status; Imputed as occupied or vacant. 3. Household Size Imputation - Occupied status confirmed but no information as to household count; the household count is imputed.
Whole-person Characteristic Imputation 4. Substitution - Population count known; all characteristics imputed for the entire household 5. Totally Allocated - Population count known; all characteristics imputed for some, but not all, persons in the household.

Note: Any housing unit imputed as occupied during count imputation will also have its household count imputed, which results in whole-person imputations.

3. Components of Census Coverage Error

This section shows important aspects of how the CCM program is implementing the concepts for the coverage error components. These implementation decisions are a trade off between the goals of the CCM program and the sample size, field work and matching resources available. Implementation of the net error concepts is similar to what we have done in the past. Some differences between net error and component error concepts are noted below.

Table 3 identifies the components of census coverage for the 2010 CCM for both people in housing units and housing units. For both universes, we will be estimating the number of correct enumerations, erroneous enumerations and omissions. Sections 3.1 to 3.4 provide more information on these coverage error components.

Table 3: Components of Census Coverage

People in Housing Units	Housing Units
1. Correct Enumerations 2. Erroneous Enumerations 3. Omissions 4. Whole-person Census Imputations	1. Correct Enumerations 2. Erroneous Enumerations 3. Omissions

Note: Erroneous enumerations will be done overall and by type of reason.

3.1. Correct Enumerations for Components

The CCM is using the following definition of being a correct enumeration when evaluating the two coverage universes in the 2010 Census:

> The enumeration is considered to be correctly enumerated if the record corresponds to a person or housing unit that should have been included anywhere in the coverage universe. If such a person or unit was included multiple times, one of the enumerations will be considered correct and the other enumerations will be erroneous.

This definition of correct enumeration for components of census coverage is different than the definition of correct enumeration used by the CCM for estimating net error. The definition for net error is stricter as it applies additional criteria of having sufficient identification information and being enumerated in the specific geographic area, referred to as block cluster search area[1].

This definition was chosen because it is similar to the definition used by the Census Bureau when estimates of erroneous enumerations and omissions were released for the 1990 and 2000 Censuses. When we determined the definition for the components for housing units, we decided to apply a similar definition for housing units as we used for people in housing units.

Based on our implementation of correct enumeration for components of census error, a person will be considered correctly enumerated by being included in a housing unit anywhere in the U.S. Additional information showing the results of whether people were enumerated in the wrong housing unit seems beneficial as well for the 2020 Census planning. The person matching and interviewing can determine if the person should have been enumerated in a housing unit somewhere else outside of the sample block cluster search area. The CCM will also identify whether the person should have been counted in the 1) same county but outside of the block cluster search area, 2) different county in the same state or 3) different state. The results will be summarized to estimate the number of resulting cases for these three geographic definitions to the national level. Results will not be provided for individual states or counties.

The field work will be limited to the sample block cluster search area. This has implications for housing units that have been mistakenly geocoded to this block cluster. The limit on field work has implications on determining if this misgeocoded census housing unit really exists somewhere else. This limit will result in these types of cases being unresolved for component error estimation. The CCM will need to develop a missing data mechanism to account for these unresolved geocoding errors in the estimation of correct and erroneous housing units.

These limitations of matching and fieldwork are more of an issue for housing unit components than people components. For people in housing units, the CCM is able to do a national computer

[1] The geographic search area is the block cluster and the one ring of surrounding census collection blocks. A block cluster is one or more contiguous collection blocks that average 30 housing units.

search as an input to the person clerical matching. Since addresses like 101 Main Street can be present in multiple towns and cities, there are geographic limits to how far you could search for duplicate addresses. Also for people in housing units, the CCM is able to ask questions during the Person Interview or the Person Followup about other places where the person may have been counted. This identification helps target other search areas.

3.2 Erroneous Enumerations for Components

For component estimation, the CCM will report the total number of erroneous enumerations. The CCM program will also estimate erroneous enumerations by type. When examining the reasons that a case can be erroneous based on the definition of correct in Section 2.2, we decided to report results by two groupings. These two groupings apply for both people in housing units and housing units.

The two groupings for reporting estimates of erroneous enumeration by type are:

- Persons or units who should not have been enumerated at all
- Duplicate person or housing unit enumerations

Table 4 shows examples of the multiple types of erroneous enumerations that are combined into one category of "should not have been enumerated at all." The reasons for this combination are because:

- Some of the reasons (e.g., should have been enumerated in group quarters, born after Census Day, died before Census Day) happen so infrequently that there is not a unique match code for these reasons. To minimize matching error, we do not want to implement procedures for infrequently occurring events. These are identified by an overall match code about being erroneously included since they are all reasons for being out-of-scope for the CCM.

- The CCM examined the keyed data from the Person Interview and Person Followup from the 2006 CCM Test. The data were analyzed to see if these infrequently occurring reasons could be identified. If the questions were answered then some of the cases could be identified. However, if these erroneous enumerations were identified based on notes provided by the interviewer then these cases were not able to be identified.

- The CCM would need to develop a missing data mechanism for the unresolved cases that supports these infrequently occurring outcomes.

- The Census Bureau has minimum sample size requirements for reporting results so some of these may have had to have been collapsed together because of small sample sizes.

- For housing units, the 2000 Housing Unit Coverage Study (HUCS) showed, based on our definition of correct, that the two main causes of erroneous housing unit enumeration were duplication and not being a housing unit (Barrett et al., 2001).

Table 4: Types of Erroneous Enumerations Grouped Together that Should Not Have Been Enumerated At All

People in Housing Units	Housing Units
Discrepant or Fictitious Person should have been enumerated in Group Quarters Born After Census Day Died Before Census Day Other Reasons	Not a Housing Unit Does Not Exist Other Reasons

For housing units, the CCM plans on limiting the matching and field work to the sample block cluster search area. The CCM will only be searching for duplicates within that area. Any estimates of erroneous enumerations due to duplication will be based on searching for duplicate housing unit addresses in the census in that limited geographic area. This will have implications on the total number of erroneous enumerations for housing units as well.

3.3 Omissions

The CCM program has been asked to estimate the total number of omissions in the two coverage universes. For people in housing units, we have been asked to estimate the number of omissions by whether or not the housing unit was included in the census. There are no specific procedures being developed to measure omissions. The estimation strategy, as it has been in the past, is to use the results of the net coverage error and the erroneous enumeration component to obtain an estimate of omissions. Section 4.2 provides more details on this estimation approach.

One group of omissions that will not be evaluated are those people or housing units removed during the census processing. Examples of these removals include a person deleted because of identified duplication in the census or housing units identified as non-existent during a field operation. Some of these removals may be in error and could lead to possible omissions in the Census. This area of evaluation was discussed during CCM planning but was dropped as the CCM program has expanded considerably leading to resource concerns. This may be an area where more work can be done for components in the 2020 evaluation.

3.4 Census Imputations

For people in housing units, we will be tallying and reporting the number of whole-person census imputations. While the goals and objectives of our program have been expanded, the CCM program has not been given the goal to evaluate the census imputation process. The CCM program can not assess whether the individual whole-person census imputations are correct or erroneous because there is no way to followup on records whose information is based solely on imputed values. Even though the CCM could determine the number of people who should have been enumerated at a particular housing unit, this would not be particularly useful to determine how census operations can be improved in the future. Since it is not in line with our goal of estimating erroneous enumerations and omissions, the CCM is reporting the total number of

whole-person imputations. This may be an area where more can be done for the 2020 coverage evaluation based on the 2010 experience.

For the housing unit universe, the CCM will be evaluating the housing units that required status imputation as to whether they were correct or erroneous. There is a history of evaluating these units as correct or erroneous as part of the dual system estimation for housing units in previous coverage evaluations. Because of this difference, it is not necessary to report housing unit status imputations separately for this universe.

4. General Estimation Approaches

This section documents the general estimation approaches for the net error and component of coverage error items listed in Section 2.

4.1 General Estimation Approach for Net Error

Like the 1990 Post-Enumeration Survey and the 2000 Accuracy and Coverage Evaluation (A.C.E.), the 2010 CCM will be evaluating net coverage error by using Dual System Estimation (DSE) 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. As part of this estimation, we will implement operations to account for missing data and reduce the sampling and nonsampling errors in our estimates. This includes an adjustment for correlation bias using Demographic Analysis results and the two group model for both the Black and Nonblack population like was done for the A.C.E. Revision II estimation. Estimates of net error will be computed based on the difference of the dual system estimate and the census count. A positive estimate shows an undercount and a negative estimate of net error shows an overcount.

$$NetError = DSE - Census$$

4.2 General Estimation Approach for Components of Coverage Error

The general estimation approach for components falls into three areas. First, the estimates of whole-person census imputations for people in housing units will be tallied from the Census files.

Second, the estimates of correct or erroneous enumerations will be design-based estimates using the matching, followup and processing results of the person and housing unit cases from the CCM sample of census enumerations. To control variance, we will implement a two-stage ratio adjustment during the estimation by taking advantage of the finite population total of census enumerations. We will also implement operations to account for missing or unresolved enumeration status and missing characteristics.

The third is the estimation of omissions. In addition to erroneous enumerations, the CCM has been tasked with generating estimates of omissions. Results of the 1950 Post-Enumeration Survey showed that trying to use the results of a post-enumeration survey to estimate the total population by estimating directly the number of omissions and erroneous enumerations in the Census resulted in a measured undercount that was well below the results produced by Demographic Analysis (U.S. Census Bureau 1960). To account for this underestimation of omissions in the total population estimates, the Census Bureau has used dual system estimation as part of the net error estimation when using survey results to estimate the total population.

Since we use dual system estimation, there is a problem with attempting to estimate omissions from those results. For practical reasons, the dual system estimation is implemented using a very restrictive definition of "correct location" as part of the definition of correct enumeration, leading to an overstatement of both erroneous enumerations and nonmatches in a way that should balance out for net error estimation, but that does not work for omissions. The only way, conceptually, to directly address this would be to try to apply matching and followup to the whole country and believe the results. Since we are going to rely on dual system estimation for estimating the true population for net error estimation, we can still use the relationship between net error, omissions and erroneous enumerations as follows.

After generating the estimate of net error and erroneous enumerations, we will generate the estimate of omissions by summing the two. The challenge is generating an accurate estimate of correct and erroneous enumerations. In this formula, the whole-person census imputations are essentially treated as all being correct enumerations. This means that the CCM estimate of omissions does not reflect any potential omissions that might be represented by a census imputation. This approach is consistent with the estimator of omissions used by the Census Bureau for previous censuses. It is noted that a range of different assumptions could be made about what percentage of the imputations is correct or erroneous. Since this could be done, this is one reason the CCM is reporting the number of these imputations so it provides a context for our estimate of omissions.

$$NetError = Omissions - ErroneousEnumerations$$
$$Omissions = NetError + ErroneousEnumerations$$

5. Proposed Estimation Domains for People in Housing Units

This section lists the proposed estimation domains for the coverage estimates of people in housing units. The estimation domains shown here are based on demographic groups, geographic areas and census operations. Tables will summarize the estimation domains which net error and component error estimates will be generated.

5.1 Estimation Domains for People in Housing Units by Demographic Groupings

The CCM will be evaluating the coverage of the census for various demographic groupings. These demographic groupings are based on race, Hispanic-Origin, tenure, age and sex. For the

coverage of American Indians, the CCM will also be producing results for this group by being classified as on or off Reservations.

Table 5 summarizes the types of coverage error statistics for the demographic groupings. The table shows that we will be generating the net error and all of the component error estimates for the race and Hispanic-Origin, tenure and age/sex groupings[2]. We will also be doing those estimates for race/Hispanic-Origin by both tenure and age/sex groupings. Whenever we are producing estimates of erroneous enumerations, we will also be producing estimates of erroneous due to duplication. These will be similar to the results in the Further Study of Person Duplication (Mule 2002). That study showed the estimates of duplication to people in other housing units by different geographic distances (block cluster search area, same county; different county same state; and different state). The CCM will also show estimates of duplication to those people also enumerated in a group quarters by type of facility.

The CCM will be generating estimates of net error by individual years of age overall and by race/Hispanic-Origin. The use of logistic regression instead of post-stratification for the dual system estimation allows the possibility of net coverage estimates by individual years of age. For component estimates, we are not planning on doing estimates by individual years of age. For component estimates of correct and erroneous, we will be generating design-based estimates using a two-stage ratio adjustment to control the variance of the estimates.

Table 5 shows that the CCM will be producing results for race and Hispanic-Origin. The CCM will be reporting these results in multiple ways. The CCM will produce race and Hispanic-Origin results using the same seven Race/Origin domains[3] that were used in the 2000 Accuracy and Coverage Evaluation. These domains assign each person to only one race/origin domain. We realize that people are interested in estimates of coverage by race. We will generate estimates, as well, by whether the race was marked alone or in combination with other races. We will produce similar estimates by whether or not the person is of Hispanic-Origin. For these estimates of race alone or in combination and Hispanic-Origin, a person may contribute to more than one estimate. We will also produce estimates by the Black and Nonblack groupings used by the Census Bureau's Demographic Analysis program. This will allow comparisons to the estimates produced by that program.

---

[2] The CCM will be using the same age/sex groupings as the A.C.E. Revision II: 0-9, 10-17, 18-29 Male, 18-29 Female, 30-49 Male, 30-49 Female, 50+ Male and 50+ Female

[3] The seven race/origin domains were American Indians on Reservation, American Indians off Reservation, Hispanic, Non-Hispanic Black, Non-Hispanic Asian, Native Hawaiian and Pacific Islander and Non-Hispanic White and Others

Table 5: Types of Coverage Error Statistics for Demographic Groupings for People in Housing Units

Demographic Grouping	Net Error	Component			
		Correct	Erroneous	Imputations	Omissions
Race and Hispanic-Origin	Yes	Yes	Yes	Yes	Yes
Tenure	Yes	Yes	Yes	Yes	Yes
Age/Sex Groupings	Yes	Yes	Yes	Yes	Yes
Individual Years of Age	Yes	No	No	No	No
Race and Hispanic-Origin by Tenure	Yes	Yes	Yes	Yes	Yes
Race and Hispanic-Origin by Age/Sex Groupings	Yes	Yes	Yes	Yes	Yes
Race and Hispanic-Origin by Individual Years of Age	Yes	No	No	No	No

Note: When erroneous enumerations are reported, the CCM will also report the types of erroneous enumerations: a) people who should not have been enumerated at all and b) duplication.

5.2 Estimation Domains for People in Housing Units by Geographic Areas

Table 6 summarizes the types of coverage error statistics for people in housing units by geographic areas. The CCM will generate net error estimates and all of the component error estimates for the Type of Enumeration Areas (TEA). Some of the TEAs may be combined for reporting purposes. The CCM will be generating estimates of components for TEAs to provide information about improving the operations for these census enumeration areas for the 2020 Census. The CCM will be only generating estimates of net error for regions, individual states and the District of Columbia. This is being done to provide net coverage information for those geographic areas, but since census operations are not done differently by state there are no plans to produce component estimates for regions or individual states. The CCM is not planning on generating net error or component error estimates for individual counties or places.

Table 6: Types of Coverage Error Statistics for Geographic Areas for People in Housing Units

Geographic Area	Net Error	Component			
		Correct	Erroneous	Imputations	Omissions
Type of Enumeration Areas	Yes	Yes	Yes	Yes	Yes
Region	Yes	No	No	No	No
Each State	Yes	No	No	No	No
Each County	No	No	No	No	No
Each Place	No	No	No	No	No

Note: Some Types of Enumeration Area may be grouped together in reporting.
When erroneous enumerations are reported, the CCM will also report the types of erroneous enumerations: a) people who should not have been enumerated at all and b) duplication.

The CCM is planning on generating some high-level estimates of coverage for the combination of census regions and race and Hispanic-Origin groupings. This would allow us to examine the different patterns of census coverage across the regions by race and ethnicity.

5.3 Estimation Domains for People in Housing Units for Census Operations

The types of coverage error estimates for census operations is different than the previous two sections. For census operations, we will only be estimating the component estimates of correct, erroneous and whole-person census imputations. These correspond to the possible outcomes of the people in housing units included in Census 2010.

We are not generating estimates of net error for census operations. We are able to use dual system estimation to generate an estimate of the true population of people in housing units or housing units. We do not have a good estimator for the true population that should have been enumerated by mail returns or during Nonresponse Followup (NRFU). Part of the dual system estimate is people who were missed by both the Census enumeration system and the independent CCM system. Since part of the total estimate was missed by both systems, several assumptions would need be made to allocate those people to the different census operations. Also, census enumeration operations like NRFU are associated only with the Census enumeration system which is one of the dual systems in our population estimation. The CCM system can determine people who should be included in housing units in Census 2010 but since it is an independent interview of the census there is difficulty in determining if the person should have been enumerated by NRFU or not.

The Census Bureau is able to do estimates by TEAs as discussed in Section 5.2 since those are geographic areas. The CCM sample can independently list and interview the housing unit in a representative sample of block clusters that are in a TEA. The CCM can independently

determine the type of enumeration area based on the Census Day address provided by the sample respondent.

Since we are not generating estimates of net error for census enumeration operations, we will also not be generating estimates of omissions by the census operations as well. The estimator for omissions shown in Section 4.2 utilizes the net error estimate to compute that.

To evaluate these components of coverage for census operations, the CCM is exploring the following options. These are some tentative examples. They are subject to change based on any future decisions to change any of the operations currently planned for Census 2010.

- Self Enumeration versus Proxy Enumeration: This will allow a comparison of the results based on the respondent who provided the information. The self enumeration can be expanded into two categories based on whether it was a mail return or whether it was an interview with an enumerator.

- Type of Form: The CCM can examine the component estimates based on the types of form used to enumerate the person. Below are some of the different ways that a person can be enumerated in the census. The CCM is determining how to take into account that a person may be enumerated by more than one method in this evaluation.

    - Return of Initial Mailing in Mailout/Mailback
    - Return of Replacement Mailing in Mailout/Mailback
    - Nonresponse Followup enumeration
    - Return from Update/Leave or Urban Update/Leave Areas
    - Enumerated in Remote Update/Enumerate or Urban Update/Enumerate Areas
    - Coverage Followup Questionnaire
    - Be Counted Form
    - Telephone Questionnaire Assistance

6. Estimation Domains for Housing Units

This section lists the proposed estimation domains for the coverage estimates of people in housing units. The estimation domains shown here are based on characteristics of the housing unit, geographic areas and census operations. Tables show the estimation domains for which net error and component error estimates will be generated. As stated earlier, imputations are not a component of census coverage for housing units as they were for people in housing units.

6.1 Estimation Domains for Housing Units by Characteristic of the Housing Unit

Table 7 summarizes the types of coverage error statistics based on the characteristics of the housing unit. These estimation domains based on housing unit characteristics are the same used in the 2000 HUCS study of housing unit net error.

Table 7: Types of Coverage Error Statistics for Housing Units by Characteristic of Housing Unit

Characteristic of Housing Unit	Net Error	Component		
		Correct	Erroneous	Omissions
Occupancy Status	Yes	Yes	Yes	Yes
Tenure of Occupied HUs	Yes	Yes	Yes	Yes
Type of Structure	Yes	Yes	Yes	Yes
Race/Origin Domain of the Householder	Yes	Yes	Yes	Yes

Note: When erroneous enumerations are reported, the CCM will also report the types of erroneous enumerations: a) units who should not have been enumerated at all and b) duplication.

6.2 Estimation Domains for Housing Units by Geographic Area

Table 8 summarizes the types of coverage error statistics by geographic areas. The choice of these estimation domains is to produce similar housing unit coverage estimates for the same geographic areas as we are producing for people in housing units. The inclusion of Metropolitan Statistical Area by Type of Enumeration Area for net error is to be consistent with estimates of housing unit coverage produced by the A.C.E.

Table 8: Types of Coverage Error Statistics for Housing Units by Geographic Areas

Geographic Areas	Net Error	Component		
		Correct	Erroneous	Omissions
Type of Enumeration Area	Yes	Yes	Yes	Yes
Metropolitan Statistical Area by Type of Enumeration Area	Yes	No	No	No
Region	Yes	No	No	No
Each State	Yes	No	No	No
Each County	No	No	No	No
Each Place	No	No	No	No

Notes: Some Types of Enumeration Area (TEA) may be grouped together in reporting.
MSA/TEA groupings in 2000 were a) Large MSA Mailout/Mailback, b) Medium MSA Mailout/Mailback, c) Small MSA and Non-MSA Mailout/Mailback and d) All Other TEAs.
When erroneous enumerations are reported, the CCM will also report the types of erroneous enumerations: a) units who should not have been enumerated at all and b) duplication.

6.3 Estimation Domains for Housing Units for Census Operations

The CCM will be taking a similar approach to the evaluation of census coverage of housing units by census operations as we did with people in housing units in Section 5.3. The CCM will be generating only estimates of census housing units that were correct or erroneous. There will be no estimates of net error or omissions.

Since we are evaluating housing unit coverage, the focus will be on the address building activities. The CCM is identifying the different operations involved with an address being included as a housing unit for the 2010 Census. These involve operations like the Local Update of Census Addresses and Address Canvassing. Addresses can also be added or deleted by the Nonresponse Followup or the several Update/Leave or Update/Enumerate operations. When doing these analyses of address building activities, the CCM will need to account for the possibility that multiple operations may be involved with the address being included in the final census.

7. References

Barrett, Diane F., Beaghen, Michael, and Smith, .Damon (2001) "ESCAP II: Census 2000 Housing Coverage Study", Executive Steering Committee For A.C.E. Policy II, Report No. 17., October 17, 2002.

Judson, D.H., J. Gregory Robinson, Kirsten West, Antonio Bruce, Bashiruddin Ahmed, and Frederick W. Hollmann (2008), "Demographic Analysis: Options in 2008 and Beyond", paper presented at the Joint Statistical Meetings, Denver, Colorado, August 2008.

Mule, T. (2002), "A.C.E. Revision II Results: Further Study of Person Duplication" DSSD A.C.E. Revision II Estimates Memorandum Series # PP-51, U.S. Census Bureau report, December 31, 2002.

Singh, R. (2003), "2010 Census Coverage Measurement - Goals and Objectives (Executive Brief)," DSSD 2010 CENSUS COVERAGE MEASUREMENT MEMORANDUM SERIES, #A-1, Washington, D.C.

Singh, R. (2005), "2010 Census Coverage Measurement - Updated Plans," DSSD 2010 CENSUS COVERAGE MEASUREMENT MEMORANDUM SERIES, #A-6, Washington, D.C.

U.S. Bureau of the Census (1960), "The Post-Enumeration Survey: 1950," Technical Paper 4, Washington, D.C.

Whitford, D. (2008), "Overview of 2010 Census Coverage Measurement Program," DSSD 2010 CENSUS COVERAGE MEASUREMENT MEMORANDUM SERIES #A-19, July 9, 2008.