



This document was prepared by and for Census Bureau staff to aid in future research and planning, but the Census Bureau is making the document publicly available in order to share the information with as wide an audience as possible. Questions about the document should be directed to Kevin Deardorff at (301) 763-6033 or [kevin.e.deardorff@census.gov](mailto:kevin.e.deardorff@census.gov)

March 18, 2013

## **2010 CENSUS PLANNING MEMORANDA SERIES**

**No. 254**

MEMORANDUM FOR      The Distribution List

From:                      Burton Reist *[signed]*  
                                 Acting Chief, Decennial Management Division

Subject:                    2010 Census Administrative Records Use for Coverage Problems  
                                 Evaluation Report

Attached is the 2010 Census Administrative Records Use for Coverage Problems Evaluation Report. The Quality Process for the 2010 Census Evaluations, Experiments, and Assessments was applied to the methodology development, specifications, software development, analysis, and documentation of the analysis and results, as necessary.

If you have questions about this report, please contact Jason Machowski at (301) 763-4173 or David Sheppard at (301) 763-9291.

Attachment

# 2010 Census Administrative Records Use for Coverage Problems Evaluation Report

U.S. Census Bureau standards and quality process procedures were applied throughout the creation of this report.

**FINAL**

Dave Sheppard, Tim Stewart, Cynthia Rothhaas, Fred Lestina, and Elizabeth Compton  
Jason Machowski  
Damon Smith

---

Decennial Statistical Studies Division  
Decennial Management Division  
Center for Administrative Records Research



This page intentionally left blank.

## Table of Contents

Executive Summary .....	viii
1. Introduction.....	1
1.1 Scope.....	1
1.2 Intended Audience .....	1
2. Background.....	1
2.1 Statistical Administrative Records System 2009 .....	2
2.2 2010 Census Coverage Followup Operation .....	2
2.3 Past Research .....	5
2.4 National Change of Address File.....	5
3. Methodology.....	5
3.1 Identification of People Potentially Overcounted During the 2010 Census Using Statistical Administrative Records System 2009.....	5
3.2 Comparison of Statistical Administrative Records System 2009 With Household Rosters Completed During Coverage Followup.....	6
3.3 National Change of Address File.....	7
3.3.1 Comparison of National Change of Address File With the 2010 Census .....	7
3.3.2 Comparison of National Change of Address File With Movers Identified by 2010 Census Coverage Followup Operation.....	8
4. Limitations .....	8
4.1 Statistical Administrative Records System 2009 Lacks Data That Are Comprehensive and Current .....	8
4.2 Research Design Missing a Field Followup Component.....	9
4.3 Data Available From National Change of Address File Insufficient For Determining Correct Residence .....	10
5. Results.....	10
5.1 Identification of People Potentially Overcounted During the 2010 Census Using Statistical Administrative Records System 2009.....	10
5.1.1 What is the total number of households with potential overcount persons? How many of these households with potential overcount persons completed an interview in the 2010 Census Coverage Followup operation?.....	11
5.1.2 What is the total number of potential overcount persons? How many of these potential overcount persons are overlapped with a coverage problem identified for the 2010 Census Coverage Followup operation?.....	12

5.1.3	What is the total count of potential overcount persons with an additional address in the administrative records database?.....	13
5.1.4	What is the total count of potential overcount persons without an additional address in the administrative records database?.....	14
5.1.5	Of the potential overcount persons with additional addresses in the administrative records database, how many of them were in Coverage Followup completed cases? .....	14
5.1.6	What is the distribution of additional addresses by living quarters type (housing units and group quarters)? .....	15
5.1.7	Conclusions.....	16
5.2	Comparison of Household Rosters Completed During 2010 Census Coverage Followup Operation and Statistical Administrative Records System 2009 .....	17
5.2.1	What is the count of households that have administrative records persons that were not listed on the Coverage Followup roster?.....	17
5.2.2	What is the count of administrative records persons that were not listed on the Coverage Followup roster? What is the count of administrative records persons that have an additional address in the administrative records database? What is the count of administrative records persons that do not have an additional address in the administrative records database?.....	18
5.2.3	For those administrative records persons that have additional addresses in the administrative records database, how many of them are enumerated at these addresses in the census (using the Census Unedited File)?.....	18
5.2.4	For Coverage Followup completed cases with different coverage problem types, do the final household rosters match with administrative records? .....	19
5.2.5	Conclusions.....	19
5.3	Analysis of the National Change of Address File.....	20
5.3.1	What is the count of National Change Of Address File records that did not match to any Decennial Response File records? For each person listed on the National Change Of Address File, where is this person counted in the 2010 Census? .....	20
5.3.2	What is the total number of Coverage Followup movers that were also in the National Change Of Address File? Where are these movers counted in the 2010 Census?.....	23
5.3.3	Conclusions.....	26
6.	Related Evaluations, Experiments, and/or Assessments.....	26
7.	Recommendations.....	27
8.	References.....	28
	Appendix: Statistical Administrative Records System 2009 - Source File Vintage.....	29

## List of Tables

Table 1: Single Years of Age and Sex: 2010 (Partial Distribution) .....	9
Table 2: Overview of Housing Unit Counts .....	11
Table 3: Overview Housing Unit Counts by Reason for Coverage Followup .....	12
Table 4: Housing Unit Counts – Distribution of Multiple Reasons.....	12
Table 5: Overview of Person Counts .....	13
Table 6: Overview of Person Counts by Reason for Coverage Followup.....	13
Table 7: Potential Overcount Persons with At Least One Additional Master Address File Identification Number in Statistical Administrative Records System 2009 .....	14
Table 8: Potential Overcount Persons with No Additional Master Address File Identification Number in Statistical Administrative Records System 2009.....	14
Table 9: Potential Overcount Persons with At Least One Additional Address in Statistical Administrative Records System 2009 for Records in Both Statistical Administrative Records System 2009 and Coverage Followup Completed Cases by Response Type.....	15
Table 10: Potential Overcount Persons with No Additional Master Address File Identification Number in Statistical Administrative Records System 2009 for Records in Both Statistical Administrative Records System 2009 and Coverage Followup Completed Cases by Response Type .....	15
Table 11: Group Quarters/Housing Unit Status of Additional Master Address File Identification Numbers on Census Unedited File .....	15
Table 12: Potential Overcount Persons with Additional Master Address File Identification Numbers on Census Unedited File by Response Type.....	16
Table 13: Coverage Followup Housing Units with Statistical Administrative Records System 2009-Only Persons.....	17
Table 14: Coverage Followup Housing Units with Statistical Administrative Records System 2009-Only Persons and At Least One Additional Address in Statistical Administrative Records System 2009.....	18
Table 15: Statistical Administrative Records System 2009-Only Persons for Coverage Followup Rosters.....	18
Table 16: Statistical Administrative Records System 2009-Only Persons with Additional Master Address File Identification Numbers in Statistical Administrative Records System 2009 Enumerated At Same Master Address File Identification Number on Census Unedited File.....	19
Table 17: Person Match between Final Coverage Followup Roster and Statistical Administrative Records System 2009 Roster .....	19
Table 18: Results from Matching National Change of Address File and Decennial Response File .....	20
Table 19: Results from Matching National Change of Address File and Decennial Response File by National Change of Address Forwarding Date.....	21
Table 20: Results from Matching National Change of Address File and Census Unedited File..	21
Table 21: Results from Matching National Change of Address File and Census Unedited File by National Change of Address Forwarding Date.....	22
Table 22: Family Moves Only - Results from Matching National Change of Address File and Census Unedited File .....	22
Table 23: Individual Moves Only – Results from Matching National Change of Address File and Census Unedited File .....	22

Table 24: Family Moves Only - Results from Matching National Change of Address File and Census Unedited File by National Change of Address Forwarding Date ..... 23

Table 25: Individual Moves Only - Results from Matching National Change of Address File and Census Unedited File by National Change of Address Forwarding Date ..... 23

Table 26: Coverage Followup Movers – Results from Matching National Change of Address File and Census Unedited File ..... 24

Table 27: Coverage Followup Movers - Results from Matching National Change of Address File and Census Unedited File by National Change of Address Forwarding Date..... 24

Table 28: Family Coverage Followup Movers Only - Results from Matching National Change of Address File and Census Unedited File ..... 25

Table 29: Individual Coverage Followup Movers Only - Results from Matching National Change of Address File and Census Unedited File..... 25

Table 30: Family Coverage Followup Movers Only - Results from Matching National Change of Address File and Census Unedited File by National Change of Address Forwarding Date ..... 25

Table 31: Individual Coverage Followup Movers Only - Results from Matching National Change of Address File and Census Unedited File by National Change of Address Forwarding Date .... 26

## List of Figures

Figure 1: Undercount Coverage Probe .....	3
Figure 2: Overcount Coverage Probe .....	4



## Executive Summary

For more than a decade, the Census Bureau has sought ways to incorporate data collected and maintained by other government and commercial entities into the data it collects and maintains for the purpose of strengthening data quality and reducing field costs. One category of resources available for this purpose is records collected and/or maintained by government agencies (federal, state, tribal, or local) or commercial entities to administer programs or to provide services. These data sources are referred to as administrative records.

This evaluation study began an investigation into whether to expand the use of administrative records during a decennial census for identification of and potential resolution of various types of coverage problems. For the 2010 Census, the Census Bureau used a model with administrative records to help identify housing units with people who were potentially missing from the count, otherwise known as undercounted persons. The research in this evaluation begins to explore the possibility of using administrative records to identify people potentially overcounted during the 2010 Census. These overcounted people, who were counted in the wrong place or at more than one place, are referred to as erroneous enumerations.

The research in this evaluation also begins to examine how administrative records might contribute to resolving potential coverage problems found during a decennial census. This aspect of the evaluation was accomplished by comparing administrative records data to results from a 2010 Census telephone follow-up interview (that is, the Coverage Followup operation) focused on resolving selected coverage problems in the United States.

The source of administrative records data analyzed in this evaluation is the Statistical Administrative Records System 2009. The Statistical Administrative Records System 2009 is a research database containing person and address data from different administrative record sources provided to the Census Bureau from other federal agencies. Most of the input files for the Statistical Administrative Records System 2009 were delivered in May or June 2009 and reflected their content on or near April 1, 2009.

Finally, this evaluation investigated the usefulness of the National Change of Address file in determining where people lived on or about Census Day (that is, April 1, 2010). The National Change of Address file contains information needed by the United States Postal Service to forward mail from one address to another for permanent and temporary moves.

### *Can using administrative records identify potentially overcounted people?*

The results from this evaluation fail to provide a conclusive answer to this research question. There is, however, evidence that administrative records can identify people counted in the census who are not included in a household that matches to the Statistical Administrative Records System 2009. There were approximately 17.2 million potentially overcounted people identified using the Statistical Administrative Records System 2009. These potentially overcounted people fell into one of the following scenarios:

- Erroneously counted in the 2010 Census at an address where they are not included in the Statistical Administrative Records System 2009,
- Correctly counted in the 2010 Census but not included in the Statistical Administrative Records System 2009 at that address because the person was not listed on one or more of the source files for the Statistical Administrative Records System 2009, or
- Actually included at the same address in Statistical Administrative Records System 2009 as well as the 2010 Census, but unable to match because of incomplete person data in one or both sources.

Two factors associated with the design of this evaluation study limited our confidence in the number of potentially overcounted people identified by the Statistical Administrative Records System 2009. The first limitation is the absence of a field follow-up component from the research design. Without the benefit of a follow-up, we cannot determine the magnitudes by which each of the above scenarios occurred. The second limitation pertains to the lack of comprehensive and current data in the Statistical Administrative Records System 2009, which likely contributed to an overstatement of the number of potentially overcounted people. The data in the Statistical Administrative Records System 2009 were at least one year old, when compared to Census Day for the 2010 Census (that is, April 1, 2010). Therefore, the Statistical Administrative Records System 2009 is missing data for people who moved or were born from the date reflected in the source files until Census Day for the 2010 Census. Moreover, data for large households can be incomplete because some administrative sources collected information for a maximum of four dependents.

Results from the 2010 Census Coverage Measurement program, an independent measure of quality for the 2010 Census, reinforced our concerns about the number of potential overcount persons identified by the Statistical Administrative Records System 2009. The 2010 Census Coverage Measurement program evaluated how well the 2010 Census counted housing units and persons living in housing units. Estimates of net coverage error for the 2010 Census (that is, undercount or overcount) and coverage error components (that is, omissions and erroneous enumerations) are formed with data collected independently from the census for a sample survey. The 2010 Census Coverage Measurement program estimated that there were 10.0 million erroneous enumerations in the 2010 Census (Mule, 2012), which makes the 17.2 million potentially overcounted people identified by the Statistical Administrative Records System 2009 seem quite high.

*For these potentially overcounted people, can administrative records be used to find an additional address?*

Some additional addresses can be found in the Statistical Administrative Records System 2009 for potentially overcounted people. Roughly seven percent of potentially overcounted people have an additional address in Statistical Administrative Records System 2009. Furthermore, about 25 percent of potentially overcounted people found in Coverage Followup completed cases had an additional address in Statistical Administrative Records System 2009.

*For all completed Coverage Followup cases, do the final household rosters match to the household rosters for these cases in administrative records? If not, analyze the differences.*

There were roughly 4.8 million housing units with completed Coverage Followup interviews. More than half of these addresses/households contained at least one person record in the Statistical Administrative Records System 2009 that was not listed on the Coverage Followup roster. For the 2010 Census, the Coverage Followup interview was considered a “gold-plated interview,” which resulted in the household roster used for the census. Any shift to use administrative records in this capacity would likely result in major differences in outcomes for these particular types of cases.

Further research should be done to determine whether Coverage Followup is ineffective at triggering respondents to list everyone who should be considered part of the household or whether data from administrative records are too unreliable to improve coverage in the census for these types of cases. Additional research should focus on household roster differences between Coverage Followup and administrative records.

*How useful is the National Change of Address file in determining the residence status of people on or about Census Day?*

While there is similarity between where people were counted for the 2010 Census and the National Change of Address file, additional study is needed to determine which source is more likely to be in error when they disagree. Moreover, to compare Census 2010 to the National Change of Address file, we had to assume that the date to begin forwarding mail to a new address was also indicative of the date that the person actually moved to the new address. To assess the usefulness of the National Change of Address file, future research should measure the occurrence of scenarios contrary to this assumption. For example, someone might obtain a new residence and request that the forwarding of his or her mail begin before any change in residence occurs.

Most people found on the National Change of Address file and the Census Unedited File were counted at the right place in reference to Census Day. Over 86 percent of the people counted at the old address ‘moved’ on or after Census Day, while about 70 percent of those who ‘moved’ before census day were correctly counted at the new address. Relatively small numbers of people were counted at both addresses in the census. We found very little difference between the rates based on whether the move was a family move or an individual move.

### *Recommendations*

The recommendations from this evaluation study are as follows:

1. **The Census Bureau should include a follow-up component for all future studies on improving coverage with administrative records.** Coverage studies without a follow-up measurement are of limited utility because we cannot know ‘truth,’ and therefore, cannot determine if the source with the correct information is administrative records or the census (or neither). If the 2010 Census Administrative Records Use for Coverage

Problems Evaluation had included a follow-up component, more meaningful conclusions and recommendations could have been developed about the accuracy of the 17.2 million potentially overcounted people identified by the Statistical Administrative Records System 2009.

2. **The Census Bureau should continue to consider pursuing the use of administrative records to assist in resolving overcoverage problems that occur in a decennial census.** The 2010 Census Administrative Records Use for Coverage Problems Evaluation was the first study on how administrative records might help to address census duplication and erroneous enumeration. While results were unimpressive regarding the use of the Statistical Administrative Records System 2009 to identify and to correct overcoverage problems, it may be beneficial to use the most current vintage of federal data available (in combination with other data sources and other information obtained from respondents) to resolve them. Moreover, the National Change of Address File is a potentially useful source for this purpose. Most people found on the National Change of Address file and the Census Unedited File were counted at the right place in reference to Census Day. Over 86 percent of the people counted at the old address ‘moved’ on or after Census Day, while about 70 percent of those who ‘moved’ before census day were correctly counted at the new address. Relatively small numbers of people were counted at both addresses in the census. Future research should include a comparison of duplicate persons identified during census processing with the National Change of Address file.
3. **The Census Bureau should pursue additional studies focusing on why some people are in the Census but not in administrative records, and why some people are in administrative records but not in the census.** A better understanding of these groups may lead to more effective use of administrative records in the future. The 2010 Census Administrative Records Use for Coverage Problems Evaluation was hampered by missing data for people who moved or were born roughly a year or less before Census Day for the 2010 Census. Moreover, data for large households were likely incomplete because some administrative sources collected information for a maximum of four dependents. Additional research should also focus on household roster differences between Coverage Followup and administrative records.

## **1. Introduction**

### **1.1 Scope**

The objective of this evaluation was to investigate an expansion of administrative records use during a decennial census for identifying and possibly resolving potentially overcounted or erroneously enumerated people. During the 2010 Census, the Census Bureau used administrative records to identify people who were missing from the count or otherwise known as undercounted people. In contrast, the research done for this evaluation explores using administrative records to identify people potentially overcounted during the 2010 Census. It also examines whether administrative records might resolve coverage problems in a decennial census instead of resorting to a census telephone follow-up interview (that is, the Coverage Followup Operation) focused on resolving selected coverage problems in the United States.

This evaluation also investigates the usefulness of the National Change of Address (NCOA) file in determining where people lived on Census Day (that is, April 1, 2010). The NCOA file contains information residents provide the United States Postal Service to forward mail from one address to another for either permanent or temporary moves.

### **1.2 Intended Audience**

The reader of this document should have a basic understanding of the overall 2010 Census process and its operations, especially the Coverage Followup (CFU) Operation<sup>1</sup>. The authors of this evaluation report intend for it to aid research, planning, and development teams to design and to implement the 2020 Census.

## **2. Background**

Each decennial census poses the difficult challenge of counting every resident of the United States only once at his or her correct location, which is the place where the person lives or sleeps most of the time. The design and implementation of the process for conducting a decennial census is always associated with some imprecision. The result can be mistakenly omitting some people from the total count or mistakenly counting others at the wrong location, more than once, or when they should be excluded from being counted at all (for example, deaths occurring before Census Day or births occurring after it).

For more than a decade, the Census Bureau has sought ways to incorporate data collected and maintained by other government or commercial entities into data collected and maintained by the Census Bureau for the purpose of strengthening data quality and reducing field costs. One category of resources available for this purpose is records collected and/or maintained by federal, state, tribal, or local government agencies or commercial entities for the purpose of administering programs or providing services. These data sources are referred to as administrative records.

---

<sup>1</sup> For more information on the CFU, please see the 2010 Census Coverage Followup Assessment Report.

## **2.1 Statistical Administrative Records System 2009**

First produced in 1999, the Statistical Administrative Records System (StARS) is a research database containing person and address data from different administrative record sources provided to the Census Bureau from other federal agencies. These sources are as follows:

- Housing and Urban Development's Public and Indian Housing Information Center File;
- Internal Revenue Service's Individual Master File and Returns Transaction File (1040 Returns);
- Internal Revenue Service's Information Returns File (1099 Returns);
- Centers for Medicare and Medicaid Service's Active Medicare Enrollment Database File;
- Indian Health Service's Patient Registration System File;
- Selective Service System's Registration File; and
- HUD Tenant Rental Agreement Certification System File.

The StARS database also incorporates data from the Social Security Administration's Numident file. Although it was not a source of records for StARS, the Census Bureau incorporated demographic data from the Numident file into the database.

The analysis done for this research project will use data from StARS 2009. The vintage information for StARS 2009 is documented in the Appendix.

## **2.2 2010 Census Coverage Followup Operation**

Coverage improvement interviews have been used during a decennial census since Census 2000. The Census Bureau conducted these interviews to determine if they should make changes to household rosters based on responses to certain questions from initial census returns.

For the 2010 Census, the Census Bureau conducted coverage improvement interviews in the Coverage Followup Operation (CFU). There were questions in the follow-up interviews to identify missed people and people counted in error. In addition, the interviews collected missing demographic data for all persons in the household, especially for those on the continuation roster in large households. Corrections to the roster were made when necessary, and data collected from these interviews were used to determine the final census response.

About 20 million housing units were in the universe that was eligible for the 2010 CFU. The 2010 CFU eligible universe consisted of responses from the following initial census returns:

- Mailout/Mailback (including Bilingual, replacement mailings, Fulfillment, and Experimental);
- Update/Leave (U/L);
- Enumerator Questionnaires; and
- Telephone Questionnaire Assistance (TQA) interviews.

All responses in the eligible universe were in the Universe Control and Management System (UCM), had a census identification number, had a Master Address File Identification (MAFID)

number, and were nonblank forms that had sufficient information for a CFU interview (i.e., there was a last name on the form and at least one valid person with a name or age)<sup>2</sup>. In addition, cases in the eligible universe had to fall into one of the following categories or sources of coverage improvement described below:

- Large Households (LHH) are cases where the respondent-provided population count was equal to or greater than the number of possible complete person records for that form type. For example, the English Mailout/Mailback return could collect complete demographic information for six household members as well as abbreviated demographic information for six additional household members. An English Mailout/Mailback return with a respondent-provided population count of six or more would be included in the CFU universe as a case with LHH as a source of coverage improvement.
- Count Discrepancies (CD) occur in one of two ways. High CDs are cases where the number of valid people listed on the form was greater than the provided population count. Low CDs are cases where the number of valid people on the form was less than the provided population count.
- Administrative Records cases are those in which at least one person was matched between an administrative record and the census return for that housing unit and at least one person was identified on the administrative record but not on the census return.
- The Undercount Coverage Probe is a household level question pertaining to additional people staying at the household who were excluded from the household population count box. See Figure 1 for the wording of this question.

**Figure 1: Undercount Coverage Probe**

2. Were there any additional people staying here April 1, 2010 that you did not include in Question 1? Mark  all that apply.

- Children, such as newborn babies or foster children
- Relatives, such as adult children, cousins, or in-laws
- Nonrelatives, such as roommates or live-in baby sitters
- People staying here temporarily
- No additional people

If respondents answered the Undercount Coverage Probe in any of the following ways, the case became part of the eligible CFU universe:

- “Children, such as newborn babies or foster children” category only;
- “Relatives, such as adult children, cousins, or in-laws” category only;
- “Nonrelatives, such as roommates or live-in babysitters” category only; or
- “People staying here temporarily” category only.

---

<sup>2</sup> For complete details of the eligible universe definitions, please see DSSD 2010 Decennial Census Memorandum Series, No.I-06, Identification of the 2010 Coverage Followup Eligible Universe and Selection Requirements.

- The Overcount Coverage Probe is a person level question indicating that a person sometimes lives or stays somewhere else. See Figure 2 for the wording of this question.

**Figure 2: Overcount Coverage Probe**

**7. Does this person sometimes live or stay somewhere else?**

No     Yes — Mark  all that apply.

<input type="checkbox"/> In college housing	<input type="checkbox"/> For child custody
<input type="checkbox"/> In the military	<input type="checkbox"/> In jail or prison
<input type="checkbox"/> At a seasonal or second residence	<input type="checkbox"/> In a nursing home
	<input type="checkbox"/> For another reason

If respondents answered the Overcount Coverage Probe(s) in any of the following ways, the case became part of the eligible CFU universe:

- “In College Housing” category only;
  - “In the Military” category only;
  - “In Jail or Prison” category only;
  - “In a Nursing Home” category only;
  - “Household Multiple,” where multiple people on one return marked different overcount categories; or
  - “Person Multiple,” where at least one person on the return marked more than one overcount category.
- Finally, some cases became part of the eligible CFU universe for purposes of evaluation studies. Even though these types of cases would have completed CFU interviews, the results from them were not considered when determining the final makeup of these households for the 2010 Census.

One component of the evaluation cases was those that had one of the following responses to the Overcount Coverage Probe:

- “For Child Custody”
- “At a Seasonal or Second Residence” or
- “For Another Reason”

Another component of the evaluation cases was for the 2010 Census Effectiveness of Unduplication Evaluation. The goal of this evaluation was to learn about the universe of people who appear in the population count more than once. Using computer-matching algorithms, the universe of all census housing unit returns was matched against itself to identify people who may have been duplicated in the 2010 Census. A sample of potentially duplicated people was sent to CFU to determine how well the operation could resolve these cases in a decennial census environment. These types of cases are referred to later in this report as “unduplication.”



## **2.3 Past Research**

One of the decennial programs that administrative records data have contributed to is the CFU Operation. Households with potential within-household undercoverage were identified using statistical models based on data from past census tests and administrative records. These models determined a probability that a household roster was missing a person. Household rosters with a relatively high probability of missing someone were selected for CFU interviews.

The 2005 National Census Test (Sheppard, et. al., 2007) provided the first test of this administrative records modeling methodology. Households that were identified by the administrative records modeling process as having a large predicted probability of missing persons were interviewed by telephone. These follow-up interviews were successful in capturing missing persons to a limited extent.

The administrative records modeling methodology was used as part of the CFU Operation for both the 2006 Census Test (Krejsa, et. al., 2007) and the 2008 Census Dress Rehearsal (Kennel and Resnick, 2008). Using administrative records modeling to identify households containing missing persons showed promise in both tests. A similar model-based process was used in the 2010 Census.

## **2.4 National Change of Address File**

The NCOA database contains data collected and maintained by the United States Postal Service (USPS). The USPS collects change of address information from the public for providing mail-forwarding services to customers who have permanent or temporary moves. These services are available to individuals, families, and businesses. The person filling out the form requests an earliest date for the USPS to begin forwarding mail to a new location.

The Change of Address (COA) form must be filled out completely and signed. Once completed, the COA form may be submitted to a local Post Office, handed to a mail carrier, or dropped into a mail collection box. COA forms may also be filed via the Internet or telephone for a one dollar verification fee. The USPS recommends that a COA form be filled out two weeks prior to moving. The earliest a COA form can be submitted prior to moving is three months.

## **3. Methodology**

### **3.1 Identification of People Potentially Overcounted During the 2010 Census Using Statistical Administrative Records System 2009**

A potentially overcounted person is someone listed on the household roster as reported on the initial census form but not found in StARS 2009. To identify people who were potentially overcounted in the 2010 Census, analysts conducted a series of matching and merging steps to create an analysis file.

The first step in creating this analysis file involved a housing-unit-level match using MAFID between the 2010 CFU eligible universe and StARS 2009 (See Section 2.2 - 2010 Census

Coverage Followup Operation for a discussion of the CFU eligible universe). These cases were chosen because they were most likely to have coverage problems in the 2010 Census.

Of those records that matched at the housing-unit-level, at least one person on both the 2010 CFU eligible record and the StARS 2009 record had to match. Match confirmation was established by agreement on variables such as name, date of birth, and sex. If none of the people matched between the two data sources, then those household-level matches were excluded from the analysis file.

For the housing unit records that remained, the next step was an examination of all person-level records. Each person associated with a MAFID in the CFU records was compared to every person record that had the same MAFID in StARS 2009. If the CFU records had one or more persons on the household roster that did not match to any StARS 2009 data within the same MAFID, then these CFU person-level records were flagged as potential overcount persons. The person-level records that matched within the same MAFID were disregarded.

To determine whether a person was potentially missed or only counted somewhere else in the 2010 Census, census person-level records that were flagged as potential overcount persons were matched to all records in StARS 2009. Match confirmation was established by agreement on variables such as name, date of birth, and sex. Any potential overcount person records that matched to any StARS 2009 person records with a different MAFID were flagged to indicate that they had an additional address in StARS 2009. Those potential overcount person records that failed to match any StARS 2009 person records with an additional MAFID were flagged as not having any additional addresses in StARS 2009.

Any census person-level records that had matched to any additional MAFIDs were matched to the CUF. This check was done to determine if they were enumerated at these additional addresses at the final census record.

### **3.2 Comparison of Statistical Administrative Records System 2009 With Household Rosters Completed During Coverage Followup**

Using the completed 2010 Census CFU household rosters, this analysis examined whether the persons listed on the CFU household rosters were also found in StARS 2009 at the same household. Household rosters from completed CFU cases were chosen because, of all of the interviews conducted for the 2010 Census, these were the most thorough. If they were not found at the same household, further searching was done to determine if they were found elsewhere in StARS 2009 or not at all.

Analysts created another file to answer the research questions for this section of the evaluation. The first step was to obtain a file of the 2010 Completed CFU cases that included both CFU added and deleted persons.

The second step was a housing-unit-level match by MAFID between CFU completed cases to StARS 2009. If there was a housing-unit-level match, the within-household person-level records were examined. Each person associated with a MAFID in StARS 2009 was compared to every

person record that has the same MAFID in the CFU completed cases file. Match confirmation was established by agreement on variables such as name, date of birth, and sex. Flags were set on this file as described below:

- Confirmed Resident - a person matches between the CFU completed case and StARS 2009 within the MAFID
- Census only - a person is in the CFU completed case and not in StARS 2009 within the MAFID (perhaps a potential census overcount)
- Administrative Records Person only - a person is in StARS 2009 and not in the CFU completed case within the MAFID (perhaps a potential census undercount)
- Non-match - no housing-unit-level match, each census response person-level record was flagged as a non-match

For any census person-level records with a census only flag, administrative records person flag, or non-match flag, a search was conducted to see if this person was found with an additional MAFID elsewhere in StARS 2009. If a person was not found anywhere else in StARS 2009 and had a census only flag, then its flag was changed to unknown person.

Any census person-level records matched with additional MAFID(s) were matched to the CUF to check if they were enumerated at these additional addresses. The analysis file indicated at which of the five additional MAFIDs they were enumerated in the census.

### **3.3 National Change of Address File**

#### *3.3.1 Comparison of National Change of Address File With the 2010 Census*

This research examined the usefulness of the National Change of Address (NCOA) file in determining the residence status of persons on or about Census Day (April 1, 2010). If there was a person match between the NCOA file and the 2010 Decennial Response File (DRF), analysts examined where these people were counted in the 2010 Census relative to their mail-forwarding dates as reported on the NCOA file.

First, analysts extracted records from the NCOA Active Database with the effective move dates of February, March, April, and May 2010. The file included variables such as Type of Move (individual/family moves only), Name, Old Address, New Address, Effective Move Date (month and year), and Type of Move (permanent or temporary). For temporary moves, there was an indication of how long the person would stay at the new address.

Analysts then matched the DRF to the NCOA file by NCOA old address and name. If there was a match, the census person-level record was flagged to indicate that it matched to the NCOA old address. When matches occurred, analysts searched for the person on the CUF at the NCOA old address. Person records were flagged when they were found on the CUF at the NCOA old address. Analysts followed the same process when they matched the DRF to the NCOA file by NCOA new address and name.

### *3.3.2 Comparison of National Change of Address File With Movers Identified by 2010 Census Coverage Followup Operation*

This research examined the usefulness of the NCOA file in determining the residence status of movers identified by CFU on or about Census Day. If there was a person match between the NCOA file and CFU movers file, analysts examined where these CFU movers are counted in the 2010 Census relative to their mail-forwarding dates as reported on the NCOA file

The process to create this analysis file is nearly identical to the description in the previous section.

First, analysts extracted records from the NCOA Active Database with the effective move dates of February, March, April, and May 2010. The file included variables such as Type of Move (individual/family moves only), Name, Old Address, New Address, Effective Move Date (month and year), and Type of Move (permanent or temporary). For temporary moves, there was an indication of how long the person would stay at the new address.

Analysts then matched the 2010 CFU Movers file to the NCOA file by NCOA old address and name. If there was a match, the census person-level record was flagged to indicate that it matched to the NCOA old address. When matches occurred, analysts searched for the person on the CUF at the NCOA old address. Person records were flagged when they were found on the CUF at the NCOA old address. Analysts followed the same process when they matched the 2010 CFU Movers file to the NCOA file by NCOA new address and name.

## **4. Limitations**

### **4.1 Statistical Administrative Records System 2009 Lacks Data That Are Comprehensive and Current**

StARS 2009 was used for this evaluation study because it reflected the most current administrative records data available at the time of the 2010 Census. It is important, however, to acknowledge an overall limitation of the administrative records data from StARS 2009: Its data were at least one year old, when compared to Census Day (that is, April 1, 2010) for the 2010 Census. Therefore, the StARS 2009 is missing data for persons who moved or were born from the time the database was created until Census Day for the 2010 Census.

Most of the input files for StARS 2009 were delivered in May or June 2009 and reflected their content on or near April 1, 2009. The final delivery of data for tax year 2008, which was delivered in installments from the Internal Revenue Service (IRS) beginning in October 2009, was received in January 2010. Even though the tax data were from tax year 2008, the data from the IRS were useful for creating StARS 2009. IRS individual income tax return data for tax year 2008 were due for most of the American public on April 15, 2009. While the income data are for 2008, the name, address, and Social Security Number (SSN) data are from the time of filing. Therefore, the name, address, and SSN data have a reference date near April 15, 2009 for the majority of Americans (Marshall and O'Hara, 2010).

After the completion of all processing and validation steps, the StARS 2009 database contained approximately 300.5 million unique persons. On the other hand, the 2010 Census reported that the resident population of the United States was 308,745,538. The difference of over eight million people guaranteed that some people counted in the 2010 Census would be missing from StARS 2009.

The 2010 Census Match Study, which included an examination of administrative records that were different but similar to StARS 2009, yielded quantitative results that help to explain the observance of more people in the 2010 Census population count than in StARS 2009 (Rastogi and O’Hara, 2012). According to the 2010 Census Match Study, there were 5.5 million 2010 Census persons<sup>3</sup> that were not found in administrative records. About 4.0 million of these persons were children under the age of 17. One reason why this age group is less likely to be in administrative records compared to the 2010 Census pertains to tax data. Tax data is one important source of information on children in administrative records. Babies born on or after January 1, 2010 would not be claimed on 2009 taxes, therefore they may be reported in the 2010 Census, but would not be in the administrative records data obtained by the Census Bureau. Additionally, tax forms such as 1040EZ do not collect data on dependents. There were also a number of dependents in administrative records that were not eligible for matching because there was not enough information to validate these records. Finally, the IRS 1040 data used in the 2010 Census Match Study only had information on a maximum of four dependents per tax return, potentially limiting the number of children reported in larger households.

The same issues identified by the 2010 Census Match Study also apply to the StARS 2009 data used for this evaluation. In fact, the limitation is greater for StARS 2009 because the files are from 2009, and the tax data are for tax year 2008. For example, babies born on or after January 1, 2009 would not appear in 2008 tax records. See Table 1 for a partial breakdown of age and sex according to the 2010 Census: There were approximately eight million persons counted in the 2010 Census who were one year old or under 1 year old.

**Table 1: Single Years of Age and Sex: 2010 (Partial Distribution)**

Age	Number		
	Both sexes	Male	Female
<b>Total population (all ages)</b>	308,745,538	151,781,326	156,964,212
<b>Under 5 years</b>	20,201,362	10,319,427	9,881,935
Under 1 year	3,944,153	2,014,276	1,929,877
1 year	3,978,070	2,030,853	1,947,217

Source: 2010 Census Summary File 1

## 4.2 Research Design Missing a Field Followup Component

Most coverage related studies are fieldwork intensive. We usually identify the accuracy of the composition of specific housing units by conducting a follow-up interview. This interview varies depending on the goals of the study. They can be conducted over the phone or in person,

<sup>3</sup> To have the potential to be matched from the 2010 Census to administrative records (and vice versa), each person has to be assigned a unique person identifier or a protected identification key (PIK). Some 2010 Census persons lacked sufficient data to uniquely identify them, so they could not be assigned a PIK. The 5.5 million persons referenced here were assigned a PIK, but not found in administrative records.

and can be structured as dependent (where the respondent is reminded of the household roster previously submitted) or independent (where a new household roster is gathered, matched to the original roster, and the discrepancies are discussed). These studies include the 2010 Census Coverage Followup Assessment, the 2010 Census Quality Survey, and several of the operations in the 2010 Census Coverage Measurement program.

This study, however, has no follow-up component. Without a follow-up component, this study will be able to observe discrepancies between the 2010 Census and StARS 2009, but cannot comment on which source is accurate.

While this study attempts to deal with this deficiency by also specifically reviewing households that completed a CFU interview, these cases will not be representative of the universe of potentially overcounted people identified using StARS 2009.

#### **4.3 Data Available From National Change of Address File Insufficient For Determining Correct Residence**

The change of address information collected by the post office allows the USPS to forward mail from an old address to a new address for a fixed period of time. This time period begins on the effective date submitted by the postal customer. To determine, however, if a person should be counted in the census at the old or the new address, we need to know the effective move date. Forwarding may be requested prior to a move (for example, if the new residence has already been obtained) or after a move (for example, if the housing unit at the old address has not yet been sold and the person still has access to the mailbox at the old address).

Old or new addresses from the Change of Address form may include Post Office Boxes and other non city style addresses, which would not match (or would unlikely match) to addresses counted in the census. In addition, the file of change of address requests has limitations in terms of content. For family moves, it contains the name of only one person -- as opposed to listing each person included in the move.

### **5. Results**

#### **5.1 Identification of People Potentially Overcounted During the 2010 Census Using Statistical Administrative Records System 2009**

Overall, the reader must note the limitations associated with the administrative record data used for this evaluation. The data used for this study were from the StARS 2009 database. StARS 2009 is a collection of administrative records sources from the year 2009. Therefore, results from this section may not contain the most updated information for persons who moved or were born from the time StARS 2009 was developed up to Census Day in 2010.

Potential overcount persons are persons listed on the household roster as reported on the initial census form, but not found in administrative records using StARS 2009. At least one person on the initial census form and in administrative records must match before labeling other persons as potential overcount persons. The reality is that these 'potentially overcounted people' may have

been erroneously included in the census; they may be missing from the StARS 2009 at that address; or they may be in fact included at that address in both the census and in StARS 2009, but we are unable to identify them as a match because of incomplete data. These results are the first look at how these data sources compare and begin to tell us where to direct our future research.

*5.1.1 What is the total number of households with potential overcount persons? How many of these households with potential overcount persons completed an interview in the 2010 Census Coverage Followup operation?*

After all processing and validation steps were completed, the StARS 2009 database contained approximately 300.5 million unique persons (Marshall and O’Hara, 2010). The resident population of the United States in the 2010 Census was 308,745,538. The difference of over eight million people will guarantee that some people in the 2010 Census will be missing from StARS 2009.

After matching the 2010 CFU eligible universe to StARS 2009, there were 11,887,559 households identified as having potential overcount persons (see Table 2). This is the number of housing unit level matches between the 2010 CFU eligible universe and StARS 2009 -- with at least one person level match and at least one person record on the census roster that is missing from the StARS 2009 roster.

The CFU had a workload of roughly 7.4 million cases (that is, cases dialed and interviews attempted). Of those cases, about 4.8 million of them or 66 percent had a completed CFU interview. For this evaluation report, these completed CFU interviews are referred to as CFU completed cases. Table 2 presents a cross tabulation of households in StARS 2009 with potentially overcounted persons by the number of CFU completed cases.

Almost all of these matches were associated with households that fell outside of the CFU completed cases. Of almost 12 million households identified as having potential overcount persons, only about 64 thousand had completed CFU interviews.

**Table 2: Overview of Housing Unit Counts**

	StARS 2009			Total
	Included	Not Included		
CFU	Included	63,908	4,737,745	4,801,653
Completed	Not Included	11,823,651	-----	11,823,651
Cases	Total	11,887,559	4,737,745	16,625,304

Source: CFU Completed Cases and StARS 2009

As expected, almost all of the 63,908 CFU completed cases that contained people identified as potential overcount persons by StARS 2009 went to CFU because of one or more overcount coverage reason. Table 3 shows that about 52 percent of these housing units went to CFU for a single overcount reason, while roughly 44 percent went for multiple reasons.

**Table 3: Overview Housing Unit Counts by Reason for Coverage Followup**

Coverage Type	Self-Response <sup>4</sup>		Enumerator Return Response <sup>5</sup>		Totals	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Overcount Total<sup>6</sup></b>	30,248	50.8%	3,041	69.8%	33,289	52.1%
- Overcount	835	2.8%	216	7.1%	1,051	3.2%
- Unduplication	29,124	96.3%	2,788	91.7%	31,912	95.9%
- CD High	289	1.0%	37	1.2%	326	1.0%
<b>Undercount<sup>7</sup></b>	1,280	2.2%	133	3.1%	1,413	2.2%
<b>Large Household</b>	1,144	1.9%	0	0.0%	1,144	1.8%
<b>Research</b>	1	<0.1%	27	0.6%	28	<0.1%
<b>Multiple Reasons</b>	26,880	45.1%	1,154	26.5%	28,034	43.9%
<b>Total</b>	<b>59,553</b>	<b>100.0%</b>	<b>4,355</b>	<b>100.0%</b>	<b>63,908</b>	<b>100.0%</b>

Source: CFU Completed Cases and StARS 2009

When further investigating the make-up of multiple reasons, we found that about 99 percent of this category included at least one overcount reason as shown in Table 4.

**Table 4: Housing Unit Counts – Distribution of Multiple Reasons**

Coverage Type	Self-Response		Enumerator Return Response		Totals	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Multiple Reasons</b>	26,880	100.0%	1,154	100.0%	28,034	100.0%
- At Least One Overcount Reason <sup>8</sup>	26,722	99.4%	1,138	98.6%	27,860	99.4%
- No Overcount Reason	158	0.6%	16	1.4%	174	0.6%

Source: CFU Completed Cases and StARS 2009

### 5.1.2 *What is the total number of potential overcount persons? How many of these potential overcount persons are overlapped with a coverage problem identified for the 2010 Census Coverage Followup operation?*

When there were housing unit level matches between StARS 2009 and the CFU universe file, StARS 2009 identified 17,209,013 potential overcount persons, as shown below in Table 5.

<sup>4</sup> Self-Response refers to Mail Out/Mail Back, Update/Leave, Puerto Rico, Fulfillment, and Experiments.

<sup>5</sup> Enumerator Return Response refers to Telephone Questionnaire Assistance, Telephone Questionnaire Assistance Experiments, and Nonresponse Followup.

<sup>6</sup> Overcount Total includes cases from the Overcount Coverage Probe, Unduplication, and Count Discrepancy High.

<sup>7</sup> Undercount includes cases from the Undercount Coverage Probe, cases identified by administrative records as having an undercount, and Count Discrepancy Low.

<sup>8</sup> At least one Overcount Reason means that the housing unit was sent to CFU because a response to the Overcount Coverage Probe, Unduplication, and/or Count Discrepancy High were present (regardless of the presence of other non-overcount reasons).



**Table 5: Overview of Person Counts**

		StARS 2009		Total
		Included	Not Included <sup>9</sup>	
CFU	<b>Included</b>	98,895	19,704,416	19,803,311
<b>Completed</b>	<b>Not Included</b>	17,110,118	-----	17,110,118
<b>Cases</b>	<b>Total</b>	17,209,013	19,704,416	36,913,429

Source: CFU Completed Cases and StARS 2009

In addition, Table 5 shows that almost all of the potential overcount persons were associated with households that were outside of the CFU production universe. Only about 98 thousand persons in households that went to CFU were identified as potential overcount persons by StARS 2009. This statistic represents only 8.0 percent of the 1,235,096 people that were deleted through CFU (Govern, 2012). Therefore, over a million of those identified as deletes in completed CFU cases were not identified as potentially overcounted persons from StARS 2009.

Furthermore, these roughly 98 thousand persons represent less than one percent of the more than 17.2 million potential overcount persons identified from the StARS 2009. Therefore, the vast majority of the potentially overcounted persons identified from StARS 2009 were not involved in completed CFU cases.

As expected, almost all of the 98,895 people in CFU completed cases who were identified as potential overcount persons by StARS 2009 were in households that went to CFU because of one or more overcount coverage reason as shown in Table 6.

**Table 6: Overview of Person Counts by Reason for Coverage Followup**

Coverage Type	Self-Response		Enumerator Return Response		Totals	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Overcount Total</b>	45,883	49.7%	4,637	70.9%	50,520	51.1%
- Overcount	1,457	3.2%	330	7.1%	1,787	3.5%
- Unduplication	43,969	95.8%	4,252	91.7%	48,221	95.5%
- CD High	457	1.0%	55	1.2%	512	1.0%
<b>Undercount</b>	1,995	2.2%	179	2.7%	2,174	2.2%
<b>Large Household</b>	3,613	3.9%	0	0.0%	3,613	3.7%
<b>Research</b>	1	<0.1%	34	0.5%	35	<0.1%
<b>Multiple Reasons</b>	40,866	44.3%	1,687	25.8%	42,553	43.0%
<b>Total</b>	<b>92,358</b>	<b>100.0%</b>	<b>6,537</b>	<b>100.0%</b>	<b>98,895</b>	<b>100.0%</b>

Source: CFU Completed Cases and StARS 2009

### 5.1.3 What is the total count of potential overcount persons with an additional address in the administrative records database?

Table 7 shows the total count of potential overcount persons with an additional address in StARS 2009 by whether the household where the person was counted was a self-response or enumerator return. Of the 17.2 million potentially overcounted persons, there were about 1.2 million persons (about 7.2 percent) with an additional address in the administrative records database. Table 7 also shows the distribution of these persons by the coverage type reason they were in CFU.

<sup>9</sup> Includes persons added in CFU.

**Table 7: Potential Overcount Persons with At Least One Additional Master Address File Identification Number in Statistical Administrative Records System 2009**

Coverage Type	Self-Response		Enumerator Return Response		Totals	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Overcount Total</b>	10,603	1.2%	1,223	0.4%	11,826	1.0%
<b>Undercount</b>	254	<0.1%	21	<0.1%	275	<0.1%
<b>Large Household</b>	372	<0.1%	0	0.0%	372	<0.1%
<b>Research</b>	0	0.0%	1	<0.1%	1	<0.1%
<b>Multiple Reasons</b>	11,697	1.3%	404	0.1%	12,101	1.0%
<b>Not in CFU</b>	898,877	97.5%	308,228	99.5%	1,207,105	98.0%
<b>Total</b>	<b>921,803</b>	<b>100.0%</b>	<b>309,877</b>	<b>100.0%</b>	<b>1,231,680</b>	<b>100.0%</b>

Source: CFU Completed Cases and StARS 2009

*5.1.4 What is the total count of potential overcount persons without an additional address in the administrative records database?*

Conversely, Table 8 shows that of the 17.2 million potentially overcounted persons, there were about 16 million persons (92.8 percent) without an additional address in StARS 2009. Those 16 million people were unable to be matched to any persons in StARS 2009. This could be because they are not in the StARS 2009, or they did not contain complete enough data in StARS 2009 and/or in the 2010 Census for us to have the ability to match them even though they were there. Table 8 also presents the distribution of these persons by whether the household where the person was counted was a self-response or enumerator return, in addition to the coverage type reason they were in CFU.

**Table 8: Potential Overcount Persons with No Additional Master Address File Identification Number in Statistical Administrative Records System 2009**

Coverage Type	Self-Response		Enumerator Return Response		Totals	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Overcount Total</b>	35,280	0.3%	3,414	0.1%	38,694	0.2%
<b>Undercount</b>	1,741	<0.1%	158	<0.1%	1,899	<0.1%
<b>Large Household</b>	3,241	<0.1%	0	0.0%	3,241	<0.1%
<b>Research</b>	1	<0.1%	33	<0.1%	34	<0.1%
<b>Multiple Reasons</b>	29,169	0.2%	1,283	<0.1%	30,452	0.2%
<b>Not in CFU</b>	12,524,632	99.5%	3,378,381	99.9%	15,903,013	99.5%
<b>Total</b>	<b>12,594,064</b>	<b>100.0%</b>	<b>3,383,269</b>	<b>100.0%</b>	<b>15,977,333</b>	<b>100.0%</b>

Source: CFU Completed Cases and StARS 2009

*5.1.5 Of the potential overcount persons with additional addresses in the administrative records database, how many of them were in Coverage Followup completed cases?*

About one quarter of potentially overcounted persons with additional addresses in StARS 2009 were in CFU completed cases. Table 9 shows that 24,575 potentially overcounted persons with an additional address in StARS 2009 were found in CFU completed cases.

**Table 9: Potential Overcount Persons with At Least One Additional Address in Statistical Administrative Records System 2009 for Records in Both Statistical Administrative Records System 2009 and Coverage Followup Completed Cases by Response Type**

Coverage Type	Self-Response		Enumerator Return Response		Totals	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Overcount Total</b>	10,603	46.3%	1,223	74.2%	11,826	48.1%
<b>Undercount</b>	254	1.1%	21	1.3%	275	1.1%
<b>Large Household</b>	372	1.6%	0	0.0%	372	1.5%
<b>Research</b>	0	0.0%	1	0.1%	1	<0.1%
<b>Multiple Reasons</b>	11,697	51.0%	404	24.5%	12,101	49.2%
<b>Total</b>	<b>22,926</b>	<b>100.0%</b>	<b>1,649</b>	<b>100.0%</b>	<b>24,575</b>	<b>100.0%</b>

Source: CFU Completed Cases and StARS 2009

On the other hand, Table 10 shows that there are 74,320 potentially overcounted persons without additional addresses in StARS 2009 appearing in CFU completed cases.

**Table 10: Potential Overcount Persons with No Additional Master Address File Identification Number in Statistical Administrative Records System 2009 for Records in Both Statistical Administrative Records System 2009 and Coverage Followup Completed Cases by Response Type**

Coverage Type	Self-Response		Enumerator Return Response		Totals	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Overcount Total</b>	35,280	50.8%	3,414	69.8%	38,694	52.1%
<b>Undercount</b>	1,741	2.5%	158	3.2%	1,899	2.6%
<b>Large Household</b>	3,241	4.7%	0	0.0%	3,241	4.4%
<b>Research</b>	1	<0.1%	33	0.7%	34	0.1%
<b>Multiple Reasons</b>	29,169	42.0%	1,283	26.3%	30,452	41.0%
<b>Total</b>	<b>69,432</b>	<b>100.0%</b>	<b>4,888</b>	<b>100.0%</b>	<b>74,320</b>	<b>100.0%</b>

Source: CFU Completed Cases and StARS 2009

### 5.1.6 What is the distribution of additional addresses by living quarters type (housing units and group quarters)?

Table 11 shows that almost all of the additional addresses identified as having persons potentially overcounted at a household in the census were housing units and not group quarters.

**Table 11: Group Quarters/Housing Unit Status of Additional Master Address File Identification Numbers on Census Unedited File**

	Number of Persons	Percent
<b>Overall</b>	<b>1,231,680</b>	<b>100.0%</b>
- Group Quarters	1,664	0.1%
- Housing Units	1,230,016	99.9%

Source: Census Unedited File

In addition, Table 12 shows that almost three quarters of the additional addresses identified for persons potentially overcounted at a household in the census were enumerated on self response forms.

**Table 12: Potential Overcount Persons with Additional Master Address File Identification Numbers on Census Unedited File by Response Type**

	<b>Number of Persons</b>	<b>Percent</b>
<b>Overall</b>	<b>1,231,680</b>	<b>100.0%</b>
- Self-Response	921,803	74.8%
- Enumerator Return Response	309,877	25.2%

Source: Census Unedited File

### 5.1.7 Conclusions

*Can using administrative records identify potentially overcounted people?*

The results from this evaluation fail to provide a conclusive answer to this research question. There is, however, evidence that administrative records can identify people counted in the census who are not included in a household that matches to the StARS 2009. There were approximately 17.2 million potentially overcounted people identified using the StARS 2009. These potentially overcounted people fell into one of the following scenarios:

- Erroneously counted in the 2010 Census at an address where they are not included in the StARS 2009,
- Correctly counted in the 2010 Census but not included in the StARS 2009 at that address because the person was not listed on one or more of the source files for the StARS 2009, or
- Actually included at the same address in StARS 2009 as well as the 2010 Census, but unable to match because of incomplete person data in one or both sources.

Without the benefit of a follow-up component of this study, we cannot determine the magnitudes by which each of these situations occurred.

Results from the 2010 Census Coverage Measurement (CCM) program, an independent measure of quality for the 2010 Census, reinforced our concerns about the number of potential overcount persons identified by the StARS 2009. The 2010 CCM program evaluated how well the 2010 Census counted housing units and persons living in housing units. Estimates of net coverage error for the 2010 Census (that is, undercount or overcount) and coverage error components (that is, omissions and erroneous enumerations) are formed with data collected independently from the census for a sample survey. The 2010 CCM program estimated that there were 10.0 million erroneous enumerations in the 2010 Census (Mule, 2012), which makes the 17.2 million potentially overcounted people identified by the StARS 2009 seem quite high.

In addition, over a million of those identified as deletes in CFU completed cases were not identified as potentially overcounted persons from StARS 2009. These would have gone unnoticed if the Census Bureau had relied solely on StARS 2009 to identify overcounted persons instead of using the methods of the 2010 Census (CFU cases selected by the use of coverage questions on initial returns as well as large households and count discrepancies).

*For these potentially overcounted people, can administrative records be used to find an additional address?*

Some additional addresses can be found in the StARS 2009 for potentially overcounted people. Roughly seven percent of potentially overcounted people have an additional address in StARS 2009. Furthermore, about 25 percent of potentially overcounted people found in CFU completed cases had an additional address in StARS 2009.

## **5.2 Comparison of Household Rosters Completed During 2010 Census Coverage Followup Operation and Statistical Administrative Records System 2009**

Once again, the reader must recall the limitations pertaining to the administrative record data used for this evaluation. The data used for this study were from the StARS 2009 database. StARS 2009 is a collection of administrative records sources from the year 2009. Therefore, results from this section may not contain the most updated information for persons who moved or were born from the time StARS 2009 was developed up to Census Day in 2010.

In addition, the reader should note that the number of CFU completed cases is slightly lower in Section 5.2 compared to Section 5.1. The difference of approximately 12,962 cases appears because the analysis in Section 5.2 excluded evaluation cases (see Section 2.1 for a description of evaluation cases), while Section 5.1 included them.

### *5.2.1 What is the count of households that have administrative records persons that were not listed on the Coverage Followup roster?*

The CFU completed cases data were comprised of about 4.8 million housing units. In Table 13, about 50.8 percent of the addresses/households in the CFU completed cases universe contained at least one person record in the StARS 2009 database that was missing from the CFU roster. For the 2010 Census, the CFU interview was considered the gold-plated interview, which typically resulted in the household roster used for the census. Any shift to use administrative records in this capacity would likely result in major differences in outcomes for these particular types of cases.

**Table 13: Coverage Followup Housing Units with Statistical Administrative Records System 2009-Only Persons**

	Total Housing Units	
	Frequency	Percent
<b>CFU HU Universe</b>	4,788,691	100.0%
- Housing Units with StARS 2009-Only Persons	2,434,478	50.8%
- Housing Units with No StARS 2009-Only Persons	2,354,213	49.2%

Source: CFU Completed Cases and StARS 2009

Of the 2.4 million HUs that contained at least one StARS 2009-only person, Table 14 shows that about 11.2 percent of these HUs contained persons that had at least one additional address in the Administrative Records data. Furthermore, these HUs represented about 5.7 percent of the total CFU completed universe.

**Table 14: Coverage Followup Housing Units with Statistical Administrative Records System 2009-Only Persons and At Least One Additional Address in Statistical Administrative Records System 2009**

	Total Housing Units	
	Frequency	Percent
<b>CFU HUs with StARS 2009-only Persons</b>	2,434,478	100.0%
- Housing Units with at least one additional address in StARS 2009	273,421	11.2%
- Housing Units with no additional address in StARS 2009	2,161,057	88.8%

Source: CFU Completed Cases and StARS 2009

*5.2.2 What is the count of administrative records persons that were not listed on the Coverage Followup roster? What is the count of administrative records persons that have an additional address in the administrative records database? What is the count of administrative records persons that do not have an additional address in the administrative records database?*

Overall, there were about 19 million person records from the CFU completed cases (see Table 5). These records were evaluated using StARS 2009 data. As a result, there were person records identified as StARS 2009-only persons that were not contained on CFU rosters. Table 15 shows there were about 5.4 million administrative record only persons identified within the CFU housing unit universe. Of those 5.4 million people, about 350 thousand (about 6.7 percent) had an additional address in StARS 2009.

**Table 15: Statistical Administrative Records System 2009-Only Persons for Coverage Followup Rosters**

	Number of Persons	
	Frequency	Percent
Total StARS 2009-Only Persons	5,365,686	100.0%
- Persons with Additional MAFIDs in StARS 2009	358,134	6.7%
- Persons without Additional MAFIDs in StARS 2009	5,007,552	93.3%

Source: CFU Completed Cases and StARS 2009

*5.2.3 For those administrative records persons that have additional addresses in the administrative records database, how many of them are enumerated at these addresses in the census (using the Census Unedited File)?*

The StARS 2009 was searched to find any additional MAFIDs/addresses associated with the StARS 2009-only person. Table 15 showed that about 6.7 percent of the persons identified as StARS 2009-only persons had at least one additional address in StARS 2009. Table 16 shows that about 28 percent of the StARS 2009-only persons with additional addresses were enumerated at the additional addresses in the final 2010 Census files. The other 72 percent were not found anywhere in the 2010 Census. Some of these people may have been omissions, while others may have been in the census, but the data were not complete enough for a match to be made.

**Table 16: Statistical Administrative Records System 2009-Only Persons with Additional Master Address File Identification Numbers in Statistical Administrative Records System 2009 Enumerated At Same Master Address File Identification Number on Census Unedited File**

	Number of Persons	
	Frequency	Percent
StARS 2009- Only Persons with Additional MAFIDs	358,134	100.0%
- Persons Found at Additional MAFID in Census	100,377	28.0%
- Persons Not Found at Additional MAFID in Census	257,757	72.0%

Source: CFU Completed Cases, Census Unedited File, and StARS 2009

*5.2.4 For Coverage Followup completed cases with different coverage problem types, do the final household rosters match with administrative records?*

Recall that each person associated with a MAFID in the CFU completed cases file was compared to every person record that has the same MAFID in StARS 2009. Match confirmation was established by agreement on variables such as name, date of birth, and sex. Each person was given one of the following classifications:

- Confirmed Resident - a person matches between the CFU completed case and StARS 2009 within the MAFID
- Census only - a person is in the CFU completed case and not in StARS 2009 within the MAFID
- Non-match - no housing-unit-level match, each census response person-level record was flagged as a non-match
- Unknown – a person originally classified as a Census-only person, but who is unable to be found in StARS 2009 at another MAFID after searching for them elsewhere in StARS 2009

Table 17 shows that for the housing units with both a CFU completed case and an administrative record listing for that unit, the majority of the people (roughly 57 percent) were confirmed residents (that is, information from both sources matched). However, over three million people (roughly 16 percent) were on the roster for the CFU completed case and were missing from the roster for the matching administrative record.

**Table 17: Person Match between Final Coverage Followup Roster and Statistical Administrative Records System 2009 Roster**

	CFU Completed Cases		
	Frequency	Percent	
<b>StARS 2009 Results</b>	Confirmed Resident	11,267,875	57.0%
	Census person only	3,027,269	15.3%
	People in non-matched households	2,395,660	12.1%
	Unknown	3,084,173	15.6%
	<b>TOTAL</b>	<b>19,774,977</b>	<b>100.0%</b>

Source: CFU Completed Cases and StARS 2009

*5.2.5 Conclusions*

There were roughly 4.8 million housing units with completed CFU completed interviews. More than half of these addresses/households contained at least one person record in StARS 2009 that

was not listed on the CFU roster. For the 2010 Census, the CFU interview was considered the gold-plated interview, which resulted in the household roster used for the census. Any shift to use administrative records in this capacity would likely result in major differences in outcomes for these particular types of cases.

Further research should be done to determine whether CFU is ineffective at triggering respondents to list everyone who should be considered part of the household or whether data from administrative records are too unreliable to improve coverage in the census for these types of cases. Additional research should focus on household roster differences between CFU and administrative records.

### 5.3 Analysis of the National Change of Address File

The reader should note an important limitation associated with the data presented by type of move (that is, a family move versus an individual move). The file of change of address requests has limitations in terms of content. For family moves, it contains the name of only one person -- as opposed to listing each person included in the move. Therefore, comparisons between family moves and individual moves should show little difference because the data available for family moves are for only one family member, which essentially makes them the same as individual moves for the analysis in this section.

#### 5.3.1 *What is the count of National Change of Address File records that did not match to any address? For each person listed on the National Change of Address File, where is this person counted in the 2010 Census?*

To produce the data in this section, the NCOA file was matched to the DRF to determine whether each person was enumerated at none, one, or both of the addresses. Then the same matching was done with the CUF to see where people were counted in the 2010 Census.

Table 18 shows that more than half of the NCOA records did not match to either address. About three percent of people were enumerated at both addresses in the census.

**Table 18: Results from Matching National Change of Address File and Decennial Response File**

	Frequency	Percent
Counted at NCOA Old Address only	2,220,500	20.2%
Counted at NCOA New Address only	2,345,864	21.4%
Counted at both NCOA New and Old Addresses	286,906	2.6%
Not counted at NCOA New or Old Addresses	6,133,491	55.8%
<b>Total</b>	<b>10,986,761</b>	<b>100.0%</b>

Source: NCOA and DRF

According to the NCOA file, the matches between NCOA and DRF were almost equally split by whether the mail forwarding date was before or after Census Day. Approximately 48 percent of matches had a mail forwarding date before Census day, while the forwarding date for 52 percent was either on or after Census Day. For each person, most were enumerated in the right place (if we assume that the mail forwarding date is the move date). Table 19 shows that almost 90 percent of the people enumerated at the old address moved on or after Census Day, while about



75 percent of those who ‘moved’ before census day were correctly enumerated at the new address.

**Table 19: Results from Matching National Change of Address File and Decennial Response File by National Change of Address Forwarding Date**

	Counted at NCOA Old Address only		Counted at NCOA New Address only		Counted at both NCOA New and Old Addresses		Not Counted at NCOA New or Old Addresses	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Overall</b>	2,220,500	100.0%	2,345,864	100.0%	286,906	100.0%	6,133,491	100.0%
- NCOA Forwarding Date Before Census Day	235,851	10.6%	<b>1,754,456</b>	<b>74.8%</b>	104,158	36.3%	3,217,941	52.5%
- NCOA Forwarding Date On or After Census Day	<b>1,984,649</b>	<b>89.4%</b>	591,408	25.2%	182,748	63.7%	2,915,550	47.5%

Source: NCOA and DRF

Table 20 shows that almost all NCOA records matched to at least one address on the CUF (less than 0.5 percent were counted at neither the new nor the old NCOA addresses). Roughly six percent of people were counted at both addresses in the census.

**Table 20: Results from Matching National Change of Address File and Census Unedited File**

	Frequency	Percent
<b>CUF Result</b>	Counted at NCOA Old Address only	1,984,352 45.7%
	Counted at NCOA New Address only	2,089,950 48.1%
	Counted at both NCOA New and Old Addresses	254,456 5.9%
	Not counted at NCOA New or Old Addresses	13,649 0.3%
	<b>Total</b>	<b>4,324,407 100.0%</b>

Source: NCOA and CUF

For matches between NCOA and CUF, the NCOA file indicated that about 43 percent had a mail forwarding date before Census Day and about 57 percent had one on or after it. For each person, most were counted in the right place (if we assume that the mail forwarding date is the move date). In Table 21, almost 90 percent of the people counted at the old address moved on or after Census Day, while about 75 percent of those who ‘moved’ before census day were correctly counted at the new address.

**Table 21: Results from Matching National Change of Address File and Census Unedited File by National Change of Address Forwarding Date**

	Counted at NCOA Old Address only		Counted at NCOA New Address only		Counted at both NCOA New and Old Addresses		Not Counted at NCOA New or Old Addresses	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Overall</b>	1,984,352	100.0%	2,089,950	100.0%	254,456	100.0%	13,649	100.0%
- NCOA Forwarding Date Before Census Day	207,556	10.5%	<b>1,566,549</b>	<b>75.0%</b>	92,969	36.5%	6,107	44.7%
- NCOA Forwarding Date On or After Census Day	<b>1,776,796</b>	<b>89.5%</b>	523,401	25.0%	161,487	63.5%	7,542	55.3%

Source: NCOA and CUF

We examined the NCOA results further by whether they were family moves or individual moves. Table 22 and Table 23 show that there is little difference between where people were counted based on whether the move was a family move or an individual move. Moreover, the results in Table 22 and Table 23 are consistent with the results in Table 20. About six percent of people were counted at both addresses in the census.

**Table 22: Family Moves Only - Results from Matching National Change of Address File and Census Unedited File**

	Frequency	Percent
<b>CUF Result</b>	Counted at NCOA Old Address only	700,311 46.3%
	Counted at NCOA New Address only	713,398 47.1%
	Counted at both NCOA New and Old Addresses	93,689 6.2%
	Not counted at NCOA New or Old Addresses	6,619 0.4%
	<b>Total</b>	<b>1,514,017 100.0%</b>

Source: NCOA and CUF

**Table 23: Individual Moves Only – Results from Matching National Change of Address File and Census Unedited File**

	Frequency	Percent
<b>CUF Result</b>	Counted at NCOA Old Address only	1,284,041 45.4%
	Counted at NCOA New Address only	1,376,552 48.7%
	Counted at both NCOA New and Old Addresses	160,767 5.7%
	Not counted at NCOA New or Old Addresses	7,030 0.3%
	<b>Total</b>	<b>2,828,390 100.0%</b>

Source: NCOA and CUF

When narrowing the matches between NCOA and CUF to family moves only, the NCOA file indicated that about 44 percent had a mail forwarding date before Census Day and about 56 percent had one on or after it. For each person, most were counted in the right place (if we assume that the mail forwarding date is the move date). In Table 24, over 86 percent of the people counted at the old address moved on or after Census Day, while about 70 percent of those who ‘moved’ before census day were correctly counted at the new address.

**Table 24: Family Moves Only - Results from Matching National Change of Address File and Census Unedited File by National Change of Address Forwarding Date**

	Counted at NCOA Old Address only		Counted at NCOA New Address only		Counted at both NCOA New and Old Addresses		Not Counted at NCOA New or Old Addresses	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Overall</b>	700,311	100.0%	713,398	100.0%	93,689	100.0%	6,619	100.0%
- NCOA Forwarding Date Before Census Day	71,370	10.2%	<b>542,905</b>	<b>76.1%</b>	34,910	37.3%	2,863	43.3%
- NCOA Forwarding Date On or After Census Day	<b>628,941</b>	<b>89.8%</b>	170,493	23.9%	58,779	62.7%	3,756	56.7%

Source: NCOA and CUF

When narrowing the matches between NCOA and CUF to individual moves only, the NCOA file indicated that about 43 percent had a mail forwarding date before Census Day and about 57 percent had one on or after it. For each person, most were counted in the right place (if we assume that the mail forwarding date is the move date). In Table 25, over 86 percent of the people counted at the old address moved on or after Census Day, while about 70 percent of those who ‘moved’ before census day were correctly counted at the new address.

**Table 25: Individual Moves Only - Results from Matching National Change of Address File and Census Unedited File by National Change of Address Forwarding Date**

	Counted at NCOA Old Address only		Counted at NCOA New Address only		Counted at both NCOA New and Old Addresses		Not Counted at NCOA New or Old Addresses	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Overall</b>	1,284,041	100.0%	1,376,552	100.0%	160,767	100.0%	7,030	100.0%
- NCOA Forwarding Date Before Census Day	136,186	10.6%	<b>1,023,644</b>	<b>74.4%</b>	58,059	36.1%	3,244	46.1%
- NCOA Forwarding Date On or After Census Day	<b>1,147,855</b>	<b>89.4%</b>	352,908	25.6%	102,708	63.9%	3,786	53.9%

Source: NCOA and CUF

### 5.3.2 What is the total number of Coverage Followup movers that were also in the National Change Of Address File? Where are these movers counted in the 2010 Census?

The data presented in this section are first restricted to the number of movers identified by the CFU Operation. The CFU Operation classified 219,325 people as movers. The CFU movers file was matched to the NCOA file. The resulting matches between the CFU movers file and the

NCOA file were then matched to the CUF to determine where people were counted in the 2010 Census.

Table 26 shows that there were 25,567 CFU movers who were also found on the NCOA file. A very small number of people were counted at both addresses in the census.

**Table 26: Coverage Followup Movers – Results from Matching National Change of Address File and Census Unedited File**

CUF Result	Frequency		Percent	
	Frequency	Percent	Frequency	Percent
Counted at NCOA Old Address only	23,537	92.1%		
Counted at NCOA New Address only	2,025	7.9%		
Counted at both NCOA New and Old Addresses	5	<0.1%		
Not counted at NCOA New or Old Addresses	0	0.0%		
<b>Total</b>	<b>25,567</b>	<b>100.0%</b>		

Source: CFU Movers File, NCOA, and CUF

For matches between NCOA and the CFU movers file, the NCOA file indicated that about 36 percent had a mail forwarding date before Census Day and about 64 percent had one on or after it. For each person, most were counted in the right place (if we assume that the mail forwarding date is the move date). In Table 27, over 67 percent of the people counted at the old address moved on or after Census Day, while about 79 percent of those who ‘moved’ before census day were correctly counted at the new address.

**Table 27: Coverage Followup Movers - Results from Matching National Change of Address File and Census Unedited File by National Change of Address Forwarding Date**

	Counted at NCOA Old Address only		Counted at NCOA New Address only		Counted at both NCOA New and Old Addresses		Not Counted at NCOA New or Old Addresses	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Overall</b>	23,537	100.0%	2,025	100.0%	5	100.0%	0	0.0%
- NCOA Forwarding Date Before Census Day	7,698	32.7%	1,603	79.2%	0	0.0%	0	0.0%
- NCOA Forwarding Date On or After Census Day	15,839	67.3%	422	20.8%	5	100.0%	0	0.0%

Source: CFU Movers File, NCOA, and CUF

We further examined the NCOA results by whether they were family moves or individual moves. Table 28 and Table 29 show that there is little difference between where people were counted based on whether the move was a family move or an individual move. Moreover, the results in Table 28 and Table 29 are consistent with the results in Table 26. A relatively small number of people were counted at both addresses in the census.

**Table 28: Family Coverage Followup Movers Only - Results from Matching National Change of Address File and Census Unedited File**

		Frequency	Percent
<b>CUF Result</b>	Counted at NCOA Old Address only	6,837	92.6%
	Counted at NCOA New Address only	548	7.4%
	Counted at both NCOA New and Old Addresses	0	0.0%
	Not counted at NCOA New or Old Addresses	0	0.0%
<b>Total</b>		<b>7,385</b>	<b>100.0%</b>

Source: CFU Movers File, NCOA, and CUF

**Table 29: Individual Coverage Followup Movers Only - Results from Matching National Change of Address File and Census Unedited File**

		Frequency	Percent
<b>CUF Result</b>	Counted at NCOA Old Address only	16,700	91.9%
	Counted at NCOA New Address only	1,477	8.1%
	Counted at both NCOA New and Old Addresses	5	<0.1%
	Not counted at NCOA New or Old Addresses	0	0.0%
<b>Total</b>		<b>18,182</b>	<b>100.0%</b>

Source: CFU Movers File, NCOA, and CUF

When narrowing the matches between NCOA and the CFU movers file to family moves only, the NCOA file indicated that about 44 percent had a mail forwarding date before Census Day and about 56 percent had one on or after it. For each person, most were counted in the right place (if we assume that the mail forwarding date is the move date). In Table 30, about 59 percent of the people counted at the old address moved on or after Census Day, while about 82 percent of those who ‘moved’ before census day were correctly counted at the new address.

**Table 30: Family Coverage Followup Movers Only - Results from Matching National Change of Address File and Census Unedited File by National Change of Address Forwarding Date**

	Counted at NCOA Old Address only		Counted at NCOA New Address only		Counted at both NCOA New and Old Addresses		Not Counted at NCOA New or Old Addresses	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Overall</b>	6,837	100.0%	548	100.0%	0	0.0%	0	0.0%
- NCOA Forwarding Date Before Census Day	2,808	41.1%	447	81.6%	0	0.0%	0	0.0%
- NCOA Forwarding Date On or After Census Day	4,029	58.9%	101	18.4%	0	0.0%	0	0.0%

Source: CFU Movers File, NCOA, and CUF

When narrowing the matches between NCOA and the CFU movers file to individual moves only, the NCOA file indicated that about 32 percent had a mail forwarding date before Census Day and about 67 percent had one on or after it. For each person, most were counted in the right place (if we assume that the mail forwarding date is the move date). In Table 31, over 70

percent of the people counted at the old address moved on or after Census Day, while about 78 percent of those who ‘moved’ before census day were correctly counted at the new address.

**Table 31: Individual Coverage Followup Movers Only - Results from Matching National Change of Address File and Census Unedited File by National Change of Address Forwarding Date**

	Counted at NCOA Old Address only		Counted at NCOA New Address only		Counted at both NCOA New and Old Addresses		Not Counted at NCOA New or Old Addresses	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Overall</b>	16,700	100.0%	1,477	100.0%	5	100.0%	0	0.0%
- NCOA Forwarding Date Before Census Day	4,890	29.3%	1,156	78.3%	0	0.0%	0	0.0%
- NCOA Forwarding Date On or After Census Day	11,810	70.7%	321	21.7%	5	100.0%	0	0.0%

Source: CFU Movers File, NCOA, and CUF

### 5.3.3 Conclusions

While there is similarity between where people were counted for the 2010 Census and the NCOA file, additional study is needed to determine which source is more likely to be in error when they disagree. Moreover, to compare Census 2010 to the NCOA file, we had to assume that the date to begin forwarding mail to a new address was also indicative of the date that the person actually moved to the new address. To assess the usefulness of the NCOA file, future research should measure the occurrence of scenarios contrary to this assumption. For example, someone might obtain a new residence and request that the forwarding of his or her mail begins before any change in residence occurs.

Most people found on the NCOA file and the Census Unedited file were counted at the right place in reference to Census Day. Over 86 percent of the people counted at the old address ‘moved’ on or after Census Day, while about 70 percent of those who ‘moved’ before census day were correctly counted at the new address. Relatively small numbers of people were counted at both addresses in the census. We found very little difference between the rates presented based on whether the move was a family or an individual move.

## 6. Related Evaluations, Experiments, and/or Assessments

Assessment for the 2010 Census Coverage Measurement Recall Bias Study

2010 Census Coverage Followup Assessment

2010 Census Effectiveness of Unduplication Evaluation Report

2010 Census Match Study

## 7. Recommendations

This study began the investigation of an expansion of administrative records use during a decennial census for the purpose of identifying and potentially resolving various types of coverage problems. Our recommendations are as follows:

- **The Census Bureau should include a follow-up component for all future studies on improving coverage with administrative records.** Coverage studies without a follow-up measurement are of limited utility because we cannot know ‘truth,’ and therefore, cannot determine if the source with the correct information is administrative records or the census (or neither). If the 2010 Census Administrative Records Use for Coverage Problems Evaluation had included a follow-up component, more meaningful conclusions and recommendations could have been developed about the accuracy of the 17.2 million potentially overcounted people identified by the Statistical Administrative Records System 2009.
- **The Census Bureau should continue to consider pursuing the use of administrative records to assist in resolving overcoverage problems that occur in a decennial census.** The 2010 Census Administrative Records Use for Coverage Problems Evaluation was the first study on how administrative records might help to address census duplication and erroneous enumeration. While the results were not very promising in regard to using Statistical Administrative Records System 2009 to identify and to correct overcoverage, it may be beneficial to use the most current vintage of federal data in combination with other data sources, as well as other information obtained from respondents during the next census to resolve overcoverage problems. Moreover, the National Change of Address File is a potentially useful source for this purpose. Most people found on the National Change of Address file and the Census Unedited file were counted at the right place in reference to Census Day. Over 86 percent of the people counted at the old address ‘moved’ on or after Census Day, while about 70 percent of those who ‘moved’ before census day were correctly counted at the new address. Relatively small numbers of people were counted at both addresses in the census. Future research should include a comparison of duplicate persons identified during census processing with the National Change of Address file.
- **The Census Bureau should pursue additional studies focusing on why some people are in the Census but not in administrative records, and why some people are in administrative records but not in the census.** A better understanding of these groups may lead to more effective use of administrative records in the future. The 2010 Census Administrative Records Use for Coverage Problems Evaluation was hampered by missing data for people who moved or were born roughly a year or less before Census Day for the 2010 Census. Moreover, data for large households were likely incomplete because some administrative sources collected information for a maximum of four dependents. Additional research should also focus on household roster differences between CFU and administrative records.

## 8. References

Govern, Kelly, Julia Coombs, and Robert Glorioso. (2012) “2010 Census Coverage Followup Assessment Report.” 2010 Census Planning Memoranda Series No. 197, U.S. Census Bureau, May 29, 2012.

Heimel, Sarah and Ryan King. (2012) “2010 Census Effectiveness of Unduplication Evaluation Report.” 2010 Census Planning Memoranda Series No. 244, U.S. Census Bureau, October 10, 2012.

Kennel, Timothy and Dean Resnick. (2008) “Using Administrative Records to Improve Within Household Coverage in the 2008 Census Dress Rehearsal,” American Statistical Association 2008 Proceedings of the Section on Survey Research Methods.

Kostanich, Martine. (2009) “Identification of the 2010 Coverage Followup Eligible Universe and Selection Requirements.” DSSD 2010 Decennial Census Memorandum Series No. I-06, April 30, 2009.

Krejsa, Elizabeth, Kyra Linse, Martine Kostanich, Sarah Heimel, Leah Marshall, Edward Banz, and Ryan King. (2007) “2006 Census Test Evaluation: Evaluation 2: Coverage Improvement,” DSSD 2006 Census Test Memorandum Series No. F-05, September 24, 2007.

Marshall, Leah and Amy O’Hara. (2010) Improving and Using the Statistical Administrative Records System, Proceedings of Statistics Canada Symposium 2010.

Mule, Tom. (2012) “2010 Census Coverage Measurement Estimation Report: Summary of Estimates of Coverage for Persons in the United States”, DSSD 2010 Census Coverage Measurement Memorandum Series No. 2010-G-01, May 22, 2012.

Rastogi, Sonya and Amy O’Hara. (2012) “2010 Census Match Study Report.” 2010 Census Planning Memoranda Series No. 247, U.S. Census Bureau, November 19, 2012.

Sheppard, Dave, Christine Davies, Sarah Heimel, Ryan King, Martine Kostanich, Kyra Linse, Leann Weyl, and Henry Woltman. (2007) “2005 National Census Test: Coverage Analysis,” DSSD 2005 Memorandum Series No. E-13, June 25, 2007.



## Appendix: Statistical Administrative Records System 2009 - Source File Vintage

Source File	Vintage	Date Received
Housing and Urban Development's Public and Indian Housing Information Center File	All records entered or updated between April 1, 2008 and April 1, 2009	July 14, 2009
Internal Revenue Service Individual Master File and Returns Transaction File (1040)	Tax year (2008): All tax year 2008 records that have completed IRS processing through week 52 of 2009	Weeks 1-39: October 8, 2009  Weeks 40-52: January 5, 2010
Internal Revenue Service Information Returns File (1099)	Tax year (2008): All tax year 2008 records that have completed IRS processing through week 41 of 2009	Weeks 1-41: October 22, 2009
Centers for Medicare and Medicaid Services Active Medicare Enrollment Database File	All records with a death date after March 31, 2008 and all persons on the file known to be alive as of April 1, 2009  File cutoff date was as close as possible to September 1, 2009	December 16, 2009
Indian Health Service Patient Registration System File	All persons on the file known to be alive as of the cutoff date  File cutoff date was as close as possible to April 1, 2009	May 15, 2009
Selective Service System's Registration File	Individuals born after April 2, 1983 and on or before April 1, 1991 (ages 18-25)  File cutoff date was as close as possible to May 5, 2009	June 9, 2009
Department of Housing and Urban Development Tenant Rental Agreement Certification System File	All records entered or updated between April 1, 2008 and April 1, 2009	June 29, 2009