



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 29, 2013

**2010 CENSUS PLANNING MEMORANDA SERIES**

**No. 255**

MEMORANDUM FOR      The Distribution List

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

Subject:                    Comparative Ethnographic Studies of Enumeration Methods and  
                                 Coverage across Race and Ethnic Groups

Attached is the Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race and Ethnic Groups 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 Laurie Schwede at (301) 763-2611 or Rodney Terry at (301) 763-5475.

Attachment

MARCH 29, 2013

# **Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race and Ethnic Groups**

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

---

LAURIE SCHWEDE AND RODNEY TERRY

Center for Survey Measurement

This page intentionally left blank.

## Table of Contents

|  |     |
|--|-----|
| Executive Summary .....  | vi  |
| 1. Introduction.....   | 1   |
| 1.1 Scope.....   | 1   |
| 1.2 Intended Audience .....  | 3   |
| 2. Background.....   | 3   |
| 2.1 History of Differential Coverage and Ethnographic Studies .....  | 3   |
| 2.2 Organization of this Report.....   | 5   |
| 3. Methodology.....  | 5   |
| 3.1 Description of 2010 Census Operations and the Questionnaire .....  | 5   |
| 3.2 Questions to be Answered .....   | 7   |
| 3.3 Site Selection .....   | 8   |
| 3.4 Ethnographer Selection.....  | 10  |
| 3.5 Training.....  | 10  |
| 3.6 Field Work .....   | 11  |
| 3.7 Records Check .....  | 12  |
| 4. Limitations .....   | 13  |
| 4.1 2010 Census Operations .....   | 14  |
| 4.2 Census Coverage Measurement.....   | 14  |
| 5. Results.....  | 15  |
| 5.1 The Census Component.....  | 15  |
| 5.1.1 Records check inconsistencies: numbers, types, and reasons.....  | 20  |
| 5.1.2 Characteristics of persons and households with record inconsistencies.....   | 25  |
| 5.1.3 Distribution of record inconsistencies across race/ethnic groups and sites .....   | 27  |
| 5.1.4 Reasons for inconsistencies: ethnographic report cross-cutting themes.....   | 35  |
| 5.1.5 Ethnographic site profiles: factors contributing to identified record inconsistencies.....   | 41  |
| 5.2 The Census Coverage Measurement Person Interview Component.....  | 51  |
| 5.2.1 Records check: numbers, types, and reasons for possible coverage errors .....  | 53  |
| 5.2.2 Characteristics of persons and households with possible coverage errors .....  | 64  |
| 5.2.3 Distribution of possible coverage errors across race/ethnic groups and sites.....  | 66  |
| 5.2.4 Reasons for inconsistencies: ethnographic report cross-cutting themes.....   | 69  |
| 5.2.5 Ethnographic site profiles: factors contributing to possible coverage error .....  | 74  |
| 6. Related Evaluations, Experiments, and/or Assessments .....  | 85  |
| 6.1 Census-related Evaluations, Experiments and Assessments .....  | 85  |
| 6.2 CCM-related Evaluations, Experiments and Assessments .....   | 86  |
| 7. Key Lessons Learned, Conclusions, and Recommendations .....   | 87  |
| 7.1 Census Component .....   | 87  |
| 7.2 The CCM Component.....   | 93  |
| 8. Acknowledgments .....   | 100 |
| 9. References in the Text .....  | 101 |
| Appendix A: A Brief Review of Coverage, Ethnographic Studies, and Changing Census<br>Bureau Operations since the 1970s by Jennifer Hunter Childs ..... | 1   |
| Appendix B: Original Study Plan Research Questions .....   | 7   |
| Appendix C: The Enumerator Questionnaire and Information Sheet for the 2010 Census<br>Nonresponse Followup and Update Enumerate Operations .....       | 8   |

## List of Tables

|  |    |
|--|----|
| Table 1: Percentage Net Undercount by Race and Hispanic Origin (National CCM) .....  | 4  |
| Table 2: Locations of Race/Ethnic Research Sites .....   | 8  |
| Table 3: Interviewers Accompanied and Interviews Observed by Race/Ethnic Site .....  | 15 |
| Table 4: Distribution of Race and Ethnicity of NRFU/UE Persons in Occupied Housing<br>Units across Race/Ethnic Sites.....                | 19 |
| Table 5: Rank Order of Consistency between Data Sources for NRFU/UE Persons .....  | 22 |
| Table 6: Team Sources of Inconsistency for NRFU/UE Persons with at Least One<br>Inconsistency across Data Sources .....                  | 23 |
| Table 7: Collapsed Team Sources of Inconsistency for NRFU/UE Persons with at Least One<br>Inconsistency across Data Sources .....        | 25 |
| Table 8: Race and Ethnicity of NRFU/UE Persons with at Least One Inconsistency across<br>Data Sources .....                              | 28 |
| Table 9: NRFU/UE Persons with Inconsistency across Data Sources across Race/Ethnic Site..  | 29 |
| Table 10: Census Day Household Type for NRFU/UE Households by Race/Ethnic Sites.....   | 30 |
| Table 11: Census Day Household Type for NRFU/UE Households with at Least One<br>Inconsistency across Race/Ethnic Sites .....             | 31 |
| Table 12: Team Sources of Inconsistency for NRFU/UE Persons with at Least One<br>Inconsistency across Data Sources .....                 | 32 |
| Table 13: Specific Team Sources of Inconsistency for Persons with at Least One<br>Inconsistency across Race/Ethnic Sites .....           | 34 |
| Table 14: CCM Interviewers Accompanied and Interviews Observed by Race/Ethnic Sites.....   | 52 |
| Table 15: Analysis Sample Distribution of Race and Ethnicity across CCM Race/Ethnic<br>Sites.....  | 56 |
| Table 16: Observed CCM Persons with Possible Coverage Error .....  | 57 |
| Table 17: Sources of Possible Coverage Error for Persons with Possible Coverage Error in<br>this Evaluation .....                        | 61 |
| Table 18: All Collapsed Team Sources of Possible Coverage Error for Persons with Possible<br>Coverage Error in Observed Sites.....       | 62 |
| Table 19: Team Sources of Possible Omissions for Persons with Possible Coverage Error in<br>Observed Sites .....                         | 63 |
| Table 20: Team Sources of Possible Overcounts/Incorrect Counts for Persons with Possible<br>Coverage Error in Observed Sites.....        | 63 |
| Table 21: Census Day Household Type for Analysis Sample CCM Households by<br>Race/Ethnic Sites in Evaluation.....                        | 65 |
| Table 22: Census Day Household Type for CCM Households with Possible Coverage Error<br>across Race/Ethnic Sites in this Evaluation ..... | 66 |
| Table 23: Frequency of Persons with Possible Coverage Error by Race/Ethnic Group in this<br>Evaluation .....                             | 67 |
| Table 24: Specific Types of Coverage Error for Persons across Race/Ethnic Sites in this<br>Evaluation .....                              | 67 |
| Table 25: Team Sources of Possible Coverage Error for Persons by Site in CCM in this<br>Evaluation .....                                 | 69 |

## **Executive Summary**

The mission of the decennial census is to count everyone living in the United States once, only once, and in the right place. Accurate counts are important because census results are used to allocate seats in the House of Representatives, to redistrict, and to distribute around \$400 billion in federal funds each year. Fulfilling this mission is daunting in a country that is growing increasingly more diverse and complex.

Despite best efforts to count everyone, the U.S. Census Bureau's own record of research shows persistent differential undercounts of some minority populations across decennial censuses, such as among African Americans and Hispanics. Prior ethnographic studies conducted during census data collection operations, but separate from those operations, have identified a range of factors affecting coverage and illuminated how and why they may affect enumeration in some populations. However, previous ethnographic studies have not included systematic observations of live interviews and respondents in an actual decennial census environment.

The present evaluation was conducted to address this gap, in addition to addressing the overall need to increase the understanding of the types and sources of these persistent miscounts in the census. In this report, the authors present findings from comparative ethnographic and partial validation research. The goal of this evaluation was to explore types and sources of possible census coverage error and identify the characteristics of households and of persons affected by them through observation of live, in-person 2010 Census interviews. We aimed to identify similarities and differences across race/ethnic groups to contribute to the discussion of why differential counts persist and to suggest improvements and new research for the next census.

To the authors' knowledge, this is the first decennial census in which researchers have been able to conduct systematic ethnographic observations of enumeration methods and coverage by observing interviewers conduct live, in-person interviews with respondents within actual data collection operations in the decennial census environment. This enabled ethnographers to document factors within the interview that may have affected coverage, as well as to debrief respondents in real time.

This is a multi-stage, mixed methods evaluation. The first stage consisted of controlled comparisons of small-scale qualitative field observation studies across nine race/ethnic groups in which ethnographers accompanied interviewers to observe and audiotape (when permitted) live interviews and conduct immediate debriefings with respondents on coverage-related questions. The ethnographer observations were an independent assessment of where persons likely should be counted, reasons for any apparent miscounts, and other important enumeration issues. The second stage of this evaluation was a small-scale partial validation check of the 2010 Census by identifying inconsistencies in whether persons identified in the observed housing units were or likely should have been counted in the ethnographer-observed housing units on Census Day, April 1, 2010.

These stages were conducted on three personal-visit operations in decennial year 2010. Two of these operations were 2010 Census operations: Update Enumerate and Nonresponse Followup. The third operation was the independent Census Coverage Measurement Person Interview

operation, part of the Census Coverage Measurement Survey, conducted some months later. The intent of observing both census and later Census Coverage Measurement operations in the same site was to facilitate the identification of wider factors in the sites over the time span of successive data collections that may have affected enumeration and coverage. The initial design specified that the same ethnographer would observe both the 2010 Census and later Census Coverage Measurement operations in the same site to document changes in the site and in different sets of respondents over time. However, that design was changed in consultation with Census Coverage Measurement statisticians to have different ethnographers in the census and Census Coverage Measurement observations in each site to avoid contaminating the independence of the Census Coverage Measurement from the census, which was a core requirement in the Census Coverage Measurement dual systems estimation methodology. The evaluation methods, procedures and research questions were the same in each of the three operations.

These observation studies were conducted in sites specially selected by the evaluation authors to focus primarily, but not exclusively, on certain race/ethnic groups: African Americans in Chicago; Alaska Natives in Kodiak Borough, Alaska; American Indians on two Southwest reservations (one in Nonresponse Followup); Asians (primarily Chinese) in the San Francisco Bay area (including Chinatown during the Nonresponse Followup operation); Native Hawaiians and Other Pacific Islanders on the Big Island of Hawaii; non-Hispanic Whites in the Kansas City, Missouri area in Jackson County; Hispanics in Dallas/Fort Worth, Texas; Middle Easterners in Wayne and Macomb Counties in Michigan (during Census Coverage Measurement only); and a final generalized site in Broward, County, Florida.

There were two sets of ethnographers per site. The first set observed 2010 Census operations in either the Update Enumerate operation in March and April or the Nonresponse Followup operation in mid- to late-May, 2010 during the height of those operations. Ethnographers for the Nonresponse Followup and Update Enumerate operations accompanied 69 enumerators and observed 363 interviews. The second set of ethnographers observed Census Coverage Measurement Person Interview operation in August, 2010 within specific pre-selected Census Coverage Measurement sample clusters in these sites. Ethnographers for Census Coverage Measurement accompanied 53 interviewers and observed 318 interviews in the Census Coverage Measurement Person Interview operation.

The second stage of this evaluation was a records check. The purpose of the records check was to partially validate where persons associated with the ethnographer-observed housing units on April 1, 2010 should be counted in the census. This was done by triangulating across multiple data sources to identify inconsistencies in who should be counted at the ethnographer-observed housing units and identify types and sources of possible coverage error. A records check inconsistency is defined as when the location of where a person associated with the observed housing unit on April 1, 2010 should be counted differs between two or more data sources. Census Bureau staff matched and compared the rosters from 1) the standard interviews; 2) the ethnographers' assessments of where each person should be counted; 3) a special final localized 2010 Census unedited dataset created around each site for this evaluation and, for the Census Coverage Measurement observations, 4) the final localized 2010 Census Coverage Measurement dataset matched to the 2010 Census records around the sample clusters. It should be noted that

searches of persons with possible coverage error were not done beyond these localized records check datasets and some of these persons who may have been omitted from the localized datasets could have been included in the 2010 Census elsewhere, while some other persons could have been duplicated elsewhere in the country. To distinguish the custom localized datasets used in this evaluation from the overall national 2010 Census Unedited Files, the custom files will be identified as “final localized 2010 Census unedited dataset” and “final localized Census Coverage Measurement (or CCM) dataset. Taking into account data from all of these sources, the research team coders made their own assessment of where each person should be counted.

## **RESULTS**

This report starts with a brief summary of the numbers of enumerators/interviewers accompanied and interviews observed and audiotaped by the ethnographers in the field observation stage. The report then focuses on the results of consistency checks in terms of the numbers of records check inconsistencies in observed Nonresponse Followup/Update Enumerate household interviews or possible coverage errors in observed Census Coverage Measurement interviews, the characteristics of persons and households affected by them, and the sources of records check inconsistency or possible coverage error. It then summarizes results from the ethnographic reports and case transcripts and summaries in two ways. First, cross-cutting themes from the ethnographic reports are identified and summarized that support the sources of records check inconsistency or possible coverage error found in the records check stage. Second, brief ethnographic profiles are presented that summarize relevant geographic, economic, demographic, and socio-cultural features that provide the context for the evaluation outcomes from each site in the two different data collections. Finally, the authors offer recommendations to improve enumeration and coverage of racial and ethnic minorities as well as among non-Hispanic Whites.

### **2010 Census Operations**

During the 2010 Census Update Enumerate and Nonresponse Followup operations, eight ethnographers accompanied 69 enumerators in eight sites, with seven sites targeted primarily to different race/ethnic populations. The ethnographers observed a total of 363 live 2010 Census interviews. They audiotaped 247 of those interviews and summarized others.

#### **What are the numbers, types, and reasons for record inconsistencies identified in the records check study?**

The analysis sample for the census operations records check consisted of 786 persons who were associated with 257 occupied housing units on April 1, 2010 and that had either a verbatim transcript of a taped interview or a summary of the observed interview with sufficient information to enable it to be included in the records check. Inmovers after April 1, 2010 were excluded because they were not associated with the observed housing unit on April 1, 2010.

In terms of consistency in where each data source determined a person should be counted, 675 of 786 persons (87 percent) were assessed as counted in the same place across all four data sources. Of the 786, 103 persons (13 percent) were classified as having at least one data source whose location of where a person should be counted was inconsistent with that of another data source.



At the housing unit level, 44 (17 percent) of 257 housing units had at least one person with an inconsistency across data sources.

Sources of inconsistency are characteristics of persons, housing units, or aspects of the overall survey process (e.g., question wording) or other wider factors that can lead to possible coverage errors in surveys and censuses. Up to three inconsistency sources were coded per person. Of 103 persons who had at least one inconsistency across data sources or assessments, team members coded 25 different inconsistency sources a total of 161 times. The most-frequent sources of inconsistency were unknowledgeable respondents and question rewording. All sources of inconsistency were also collapsed into eight broad categories. Of these broader categories, interviewer error accounted for 29 percent of the sources listed, unknowledgeable respondents for 12 percent of sources listed, and respondent concealment/refusal for 9 percent for Nonresponse Followup/Update Enumerate persons with at least one inconsistency, with other sources accounting for the remaining inconsistencies. A few persons may have been missed in “hidden housing units” on the same property.

### **What are the characteristics of persons and households with record inconsistencies?**

The 18-29 age cohort was the largest cohort to have at least one mismatch across data sources. The 0-4 age cohort had the greatest proportion of inconsistency among all age cohorts.

Besides the reference person, the most frequent relationships to Person 1 with inconsistencies were biological sons and daughters. Grandchildren and nieces or nephews also had inconsistencies.<sup>1</sup> The relationship types with the greatest proportions of persons with inconsistencies within their own types were nieces/nephews (39 percent), grandchildren (33 percent), and stepsons/daughters (25 percent), and children for whom it could not be determined whether they were biological, adopted or stepchildren (23 percent).

Households were also classified into the complex/noncomplex household typology whereby a nuclear family consisting of married parents and their joint biological children, a single parent and his/her biological children, and a person living alone as noncomplex. Complex households comprise all other combinations of two or more persons in family households or nonfamily households. For example, complex family households include blended families with stepchildren; grandparents with grandchildren, and other linearly extended households as well as laterally extended households with adult siblings, aunts, cousins, nieces, and so on. They may also include one or more nonrelatives, such as an unmarried partner or a roommate, or more than one family. Complex nonfamily households are those with two or more persons, none of whom is related to the householder.<sup>2</sup>

---

<sup>1</sup> Nieces and nephews cannot be tabulated from 2010 Census relationship data because they fell into the general “other relative” category and could not be disaggregated. In this evaluation, the ethnographers identified a wide range of relationships from the standard interview or a debriefing. The numbers of relatives identified nieces/nephews, grandchildren and stepchildren were low.

<sup>2</sup> Schwede, L. (2008b), “Changes in the Distribution of Complex and Noncomplex Households in the United States: 1990 and 2000 Censuses.” Unpublished research typology for classifying household types in this evaluation. Suitland, MD: U.S. Census Bureau.

Of the 44 households that had at least one person with at least one inconsistency across data sources and assessments, 19 households were noncomplex and 17 were complex, while eight with missing data were excluded. In comparison, 162 households in the overall analysis sample were noncomplex and 82 were complex. While there appeared to be a trend toward complex household association with inconsistencies, the chi-square test of independence was not statistically significant at the designated  $p < .05$  level.

### **How are record inconsistencies distributed across the race/ethnic groups and sites?**

An analysis of inconsistent persons pooled across all sites found that among all race groups, Blacks and American Indian/Alaska Natives had the largest proportions of persons with an inconsistency. Chi-square tests of independence that individually compared persons of different race groups with one another indicated that American Indian/Alaska Natives<sup>3</sup> (one race alone) and African Americans (one race alone) were associated with inconsistency between data sources. This does not imply that there is anything inherent in this race group that causes higher possible error. Rather, this finding is associated with other variables that may cause errors.

Of the race/ethnic sites, the African American site had the most persons with at least one inconsistency, followed by the non-Hispanic White site and American Indian Site 1. Chi-square tests showed that the African American site and the first American Indian site were associated with inconsistency between data sources or assessments.

### **What reasons for record inconsistencies are found in themes that cut across the ethnographers' reports?**

For the census operations, the ethnographic reports shed light on the possible sources for records check inconsistency separately from the records check study outlined above; the focus here moves to those sources of records check inconsistency that are specifically associated with records check inconsistency. Comparing the results from the ethnographic observations, debriefings, and reports several crosscutting themes were identified that parallel the sources of error mentioned above: 1) enumerator error; 2) difficulty gaining access to respondents; 3) language issues; 4) cultural differences; and 5) mobility. A few possible "hidden households" were found.

### **What characteristics specific to each ethnographic site (documented in ethnographic profiles) contribute to the identified record inconsistencies?**

In the Alaska Native site, interviewer error, mobility for work, cycling between households, and one person's attempts to avoid the census were factors in record inconsistencies and in possibly miscounting persons. For the American Indian site, interviewer error and a lack of street names and numbers appeared to lead to possible enumeration errors. At the Hispanic site, question

---

<sup>3</sup> American Indians and Alaska Natives were grouped in one racial category in the official 2010 Census records in the dataset used. The fact that at the site level, only the American Indian site in the southwest was significantly associated with a higher number of inconsistencies indicates that it is mostly American Indians who had inconsistencies.

wording, household complexity, unknowledgeable respondents, and respondent concealment of information were factors associated with record inconsistencies.

In the Asian site, it appears that vague addresses, language barriers and different spelling of Chinese people's names in English and Chinese, listing problems, use of multiple rooms as a single housing unit, possible slow processing of forms in languages other than English, apartment building effects (when respondents pick up and fill out a questionnaire actually barcoded for another apartment or condominium in the same building, resulting in the person being counted in the wrong unit) and other issues may have been factors in record inconsistencies. In the African American site, enumerator error, mobility, possibly missed housing units, respondent concealment and hostility, mistrust of government, and confidentiality concerns are among the factors in this site. In the Non-Hispanic White site, mobility in the form of children cycling between housing units and an unknowledgeable proxy who may have provided wrong names accounted for inconsistencies across records. At the Native Hawaiian site, question wording, mobility, vague addresses, and hidden housing units were found to be sources of inconsistencies. In the Generalized site, there was just one inconsistency associated with the observed housing unit, but there were several inconsistencies caused by enumerator error in not taking the additional required steps of asking in-movers and persons in "whole-household-usual-home-elsewhere" situations if they had been included on a form for their previous April 1, 2010 residence and completing a new form if they had not or did not know.

### **The Census Coverage Measurement Person Interview Operation**

During mid- to late-August, 2010, the second set of nine ethnographers accompanied 53 interviewers in nine sites, with eight targeted primarily to different race/ethnic populations. The ethnographers observe 318 interviews. They audiotaped 252 of those interviews and summarized others.

### **What are the numbers, types, and reasons for possible coverage errors identified in the records check study?**

The analysis sample for the Census Coverage Measurement records check study consisted of 953 persons in 288 housing units in which interviews were observed who: 1) likely should have been counted at the observed housing unit on Census Day and appeared to have been correctly counted there in the final dataset; 2) likely should have been counted at the observed housing unit on Census Day but appeared to have been omitted; or 3) likely should not have been counted at the observed housing unit on Census Day but appeared to have been incorrectly counted there. The analysis sample did not include 213 persons identified in observed housing units who should have been counted elsewhere and appear to have been correctly not counted at the observed housing unit. In-movers were among those omitted.

The records check found that of 953 persons in the analysis sample, 873 (94 percent) did not have an indication of any type of possible coverage error in the localized final unedited 2010 Census dataset created for this evaluation and appeared correctly enumerated at the observed housing unit. Of the 953 persons, 60 individuals (6 percent) in 41 (14 percent) of 289 households had an indication of some type of possible coverage error, including people who were likely

omitted or counted in the wrong place, or were out of scope. Of the 60 persons with possible coverage error, thirty-two persons had possible omissions at the observed housing unit, 24 had possible incorrect counts, four were possible overcounts, meaning that they were out of scope for the census, 11 were unresolved, and nine had missing values.

For the 60 persons with an indication of possible coverage error, team members coded 24 different sources of coverage error a total of 91 times, allowing up to three sources for each person. After collapsing these sources of error into eight categories, the most-frequent sources of coverage error at the collapsed level were mobility/tenuousness, respondent confusion, and interviewer error; “hidden housing unit” was identified a few times.

### **What are the characteristics of persons and households with possible coverage errors?**

Most persons with possible coverage error were brought on to the roster with the initial roster question and had one alternative address. In terms of relationship type to the reference person, the reference person’s biological son or daughter had the most numerical instances of possible coverage error, although the aunt/uncle, roomer/boarder, and housemate/roommate were the relationship types that had the greatest proportions of possible coverage error.<sup>4</sup> Several prior logistic regression studies from the 1990 Census Post-Enumeration Survey and the Census 2000 Accuracy and Coverage Evaluation showed that distant relatives and nonrelatives were less likely to be included in the census than Person One.<sup>5</sup> Complex households can include distant relatives or nonrelatives.

Across all sites, the 0 - 4 age cohort and 18 - 29 age cohort had the greatest proportions of persons within their own cohorts with possible coverage error. A chi-square test of independence also found that instances of possible coverage error were associated with age cohorts.

Of 289 households, 41 households had at least one person with possible coverage error. Half of these 41 households had five persons or more. Households with five persons or more were significantly more likely to have possible coverage error when those other households are placed in a single group. Of these households with possible coverage errors, 18 were noncomplex and 22 were complex.<sup>6</sup> A chi-square test of independence found that possible coverage error was associated with household type, and that there were more instances of possible coverage error in complex households than in noncomplex households.

### **How are possible coverage errors distributed across the race/ethnic groups and sites?**

Of 60 persons judged to exhibit some form of possible coverage error, the race/ethnic groups with the greatest proportions of persons with possible coverage error were Native Hawaiians and Other Pacific Islanders, some other race, Hispanics of any race, and persons of multiple race. To

---

<sup>4</sup> Aunts and uncles cannot be tabulated from 2010 Census relationship data because they fell into the general “other relative” category and could not be disaggregated. In this evaluation, the ethnographers identified a wide range of relationships either from the standard interview or during a debriefing.

<sup>5</sup> Bell, W. (2004), “Distant relatives have lower census inclusion probabilities than immediate family.” Email message to David Whitford: Chief of Decennial Statistical Studies Division. May 27, 2004.

<sup>6</sup> One unit lacked information to classify by household complexity.

test significance, individual chi-square tests of independence were conducted to compare the proportion of people in a race/ethnic group with possible coverage error with all other persons placed in one group. In these analyses, Hispanics of all races were in the Hispanic group, and all race groups included persons who were Hispanic and non-Hispanic of that race.

The only race/ethnic group significantly more likely to have coverage error than the other race groups was Native Hawaiian/Other Pacific Islanders, of whom 25 percent had possible coverage error. This does not imply that there is anything inherent in this race group that causes higher possible error. Rather, this finding is associated with other variables that may cause errors.

Furthermore, persons with possible coverage error were found in the highest proportion at the Native Hawaiian site. A chi-square test of independence also found an overall association between persons in the Native Hawaiian site and instances of possible coverage error.

### **What reasons for possible coverage errors are found in themes that cut across the ethnographers' reports?**

For the Census Coverage Measurement Person Interview operation, the ethnographic reports shed light on the possible sources for enumeration error separately from the records check study outlined above; the focus here moves to those sources of enumeration errors specifically associated with possible coverage error. Comparing the results from the ethnographic observations, debriefings, and reports seven crosscutting themes were identified that parallel the sources of error mentioned above: 1) interviewer error; 2) difficulty in gaining access to respondents; 3) language barriers; 4) cultural issues; 5) mobility; 6) recall error and 7) respondent reactions to multiple cold-call visits from Census Bureau interviewers.

### **What characteristics specific to each ethnographic site (documented in ethnographic profiles) contribute to the identified possible coverage errors?**

At the Alaska Native Site, mobility for work and possible housing unit/Group Quarters operation overlap was a factor in possible coverage error as was cycling and interviewer error and a hidden housing unit that appeared to be missed by the census and Census Coverage Measurement program (Census Coverage Measurement estimates the numbers of these omissions). Several factors were contributors at the American Indian Site, including mobility for work off the reservation, general tenuousness and cycling among places, lack of street names and numbers, and remote residences. At the African American site, the ethnographer identified distrust leading to deliberate concealment of information, interviewer error, and mobility among the poor.

The ethnographer for the Native Hawaiian site documented mobility, persons with multiple residences, interviewer question rewording, and an interviewer not paying attention to what the respondent said as barriers to enumeration. At the Hispanic site, the ethnographer identified language barriers and distrust of the government and suspicion of the Census Coverage Measurement interviewers. As in the Hispanic site, the Middle Eastern and Asian sites had language barriers that prevented English-only interviewers from conducting interviews with respondents speaking Arabic and Chinese, respectively, while the Asian site also had a mismatch of the respondents' concept of who belonged in his house with the census concept of residence

based on where one usually stays and was on April 1, 2010. At the non-Hispanic White site, the ethnographer found interviewer error and lack of respondent knowledge as sources of possible coverage error. The ethnographer for the Generalized site found mobility, a hidden housing unit, lack of respondent knowledge, census form processing errors, and respondent irritation from multiple visits as reasons for possible coverage errors.

## **RECOMMENDATIONS FOR CENSUS AND CENSUS COVERAGE MEASUREMENT OPERATIONS**

Looking across the ethnographic reports and observed cases for the census and Census Coverage Measurement operations, the authors found six cross-cutting themes that could lead to coverage problems in these groups: 1) interviewer error; 2) difficulty in gaining access to respondents; 3) language barriers; 4) cultural differences; 5) mobility; and 6) recall error. The Census Coverage Measurement operation also had the additional factor of some respondents' reactions of frustration at multiple cold-call interviewer visits in both Nonresponse Followup/Update Enumerate and Census Coverage Measurement operations.

- 1) The authors offer recommendations for improving enumeration and coverage across race/ethnic groups in five areas: 1) reduce interviewer error and increase cultural awareness of minorities; 2) review, revise and test questions in the Enumerator Questionnaire and the Census Coverage Measurement Person Interview instrument; 3) translate questionnaire materials into more languages and interview those who do not speak English well in their own languages, when possible; 4) improve ways to gain access to respondents; and 5) conduct new research and consider amending the 2010 Census residence rule and situations document.

### **2) Reduce interviewer error and increase cultural awareness**

Frequent major changes to questions documented in the behavior coding study and in this section are a concern because they may affect the accuracy, completeness, validity, reliability and comparability of census data. The authors offer six suggestions for reducing interviewer error during interviews:

- **Create a new training module** that explains how questions are developed and tested and why it is important for interviewers to read questions as worded. If interviewers understand the reasons behind training them to read questions as written are related to accuracy, validity and comparability of census data in their communities, they may be more likely to read questions as written. A consistent reading of key terms, particularly the inclusion of "on April 1, 2010," when it is part of the question, is critical.
- **Develop role-playing scripts and, in training, conduct role-playing of difficult situations**, such as those with reluctant, impatient, or hostile respondents, or those that involve difficult living situations. Role-playing scenarios could be developed from some of the actual transcripts, summaries, and outcomes in this evaluation.

- **Consider reviewing and possibly rebalancing the relative importance of quantity of interviews completed compared to the quality of them in the incentive structure for rewarding and rehiring interviewers**, based on findings here that a few interviewers appear to be purposely cutting corners in order to shorten interviews and complete more interviews per day.
- **Conduct more research on the dynamic effects of respondent reluctance, anger, and resistance on interviewer behavior over the course of the interview.**
- **Develop new cultural awareness and sensitivity modules** that identify racial and ethnic minorities in the local interviewing area and that advise on appropriate behavior and issues that may arise in interviewing them.
- **Monitor interviewer behavior** beyond their initial trainee observations through increased crew leader observations. A possible new option to explore would be Computer Audio-Recorded Interviews (also known as CARI) in which interviews would be recorded and randomly monitored.

### 3) **Review, revise and test questions**

An additional important factor contributing to interviewer error is questions that are long or difficult to answer. In the Enumerator Questionnaire used in the Nonresponse Followup/Update Enumerate operations, coverage questions S2, S3, S5, roster question 1, overcount flag question 7 and omission flag question H1 were frequently misread, and the critical reference date of April 1, 2010 was often not mentioned.<sup>7</sup> For example, reduce the burden of overcount flag question 7 and omission flag question H1 by reformatting and testing these questions as separate or at least smaller chunks of questions for each of these questions' response categories. Breaking longer questions into shorter questions may also reduce respondent recall error.

In the Census Coverage Measurement Person Interview instrument, the roster question had some of the same problems, as well as the lead-in to the questions about persons staying at other places and the Group Quarters question. Some of the special paths in the instrument for alternative addresses of persons away at sea or in jail appear to need some revisions and it would appear to be useful to get more of a sense of the extent that in-movers can provide information about out-movers before ending the interview as is currently done.

### 4) **Reduce barriers to accessing respondents**

Observations in the Asian, African American, American Indian, Native Hawaiian, and Alaska Native sites found extra challenges in enumerating and interviewing rural and urban areas where finding and accessing some respondents is more difficult. Particularly for rural areas found in the American Indian and Native Hawaiian sites, the authors suggest that improved maps and additional resources, such as four-wheel drive vehicles and cell phone reimbursements may be needed to reach the remote populations in such areas to avoid undercounts. Cell phones or smart

---

<sup>7</sup> These questions can be seen in the reproduction of the Enumerator Questionnaire in Appendix C.

phones might also help in urban areas with locked gates and buzzer boxes. The lack of additional resources appeared to be a factor in the American Indian site that might have potentially affected coverage if enumerators were not able to get to all of the houses in remote areas of the reservation before the Update Enumerate operation ended.

**5) Translate questionnaire materials into more languages and interview those who do not speak English well in their own languages, when possible**

- **Hire more bilingual enumerators/interviewers**

The first suggestion would be to train and hire more bilingual interviewers so that interviews could be done on the first visit, rather than requiring followup visits or phone calls with a bilingual interviewer. The authors also suggest exploring the possible use of Audio-Computer-Assisted Self-Interviewing (also known as ACASI) for allowing respondents to be able to take the instrument from the interviewer, put on headphones, and listen and respond to audio-recorded interviews in their own languages already loaded onto the device. This would involve translating the materials into more languages, loading them into the Audio-Computer-Assisted Self-Interviewing System, and testing the instruments with respondents who do not speak English well.

- **Translate materials into more languages and improve the language card**

To support these options, the Census Coverage Measurement instrument and its supporting documents – the letter, information sheet, notice of visit, and so on – should also be translated into languages beyond the current English and Spanish to include foreign languages most frequently spoken in the U.S. These may change over time, as Asians now appear to be the fastest growing group in the United States. Findings from the Chinese site in both the census and Census Coverage Measurement components shows that some immigrant groups may need translations in more than one language, such as Mandarin and Cantonese, and that matching interviewers by both language and dialect is important in some sites. Consideration to translating the instrument into some American Indian and other indigenous languages is also suggested.

The Language Identification Card should be adapted to distinguish whether Chinese persons speak Mandarin or Cantonese so that the interviewer and crew leaders can assign an appropriate bilingual Chinese-speaking interviewer to conduct the interview.

**6) Conduct new research and consider amending the residence rule and situations document**

- **Mobility**

The authors suggest research to improve the enumeration of mobile persons at risk of coverage error, such as college students and children who are moving between and among households which are not always under formal joint custody arrangements and others who may be moving between housing units or group quarters. The authors especially suggest research to better account for the living situations of persons living seasonally on commercial fishing vessels.



Consider amending the residence rule and situations document to instruct where such persons should be counted. Currently, this living situation is not directly covered in the key residence rule and residence situations document for the 2010 Census (Lamas, 2009).<sup>8</sup>

- **Race, Hispanic Origin, Household Structure and Coverage as the United States Diversifies**

The 2010 Census shows that the country is rapidly diversifying, with almost all racial and ethnic groups growing faster than non-Hispanic Whites,<sup>9</sup> partly due to increasing immigration from Asia, Central and South America, and elsewhere. Other research indicates that minorities are more likely to live in some types of complex households than non-Hispanic Whites (Schwede, 2007)<sup>10</sup> and prior census results indicated that complex households increased from 18 percent in the 1990 Census to 21 percent in Census 2000 (Schwede 2008b). The comparable research has not yet been done to learn if that trend continued over the decade leading to the 2010 Census; though other research shows that complex family households in general increased during the 2007-2009 years of the recession and were associated with unemployment.<sup>11</sup> Additionally, research on young adults indicated that the proportion of adult children older than college students aged 25 to 34 living with their parents increased from 22 percent in 2010 to 29 percent a year or so later<sup>12</sup> and that those aged 25 to 29 increased their residential dependency and may have fared worse than others from the recession.<sup>13</sup> Some of the effects of the recession in terms of higher unemployment and higher foreclosure rates lingered through the 2010 Census data collections.

Census Bureau projections indicate that the country will reach majority minority status sometime in the 2040s. Challenges to maintaining or improving coverage may increase in tandem with the growth of persistently hard-to-count groups and the possible growth of complex households. Additional research on the interrelationships of race, Hispanic origin, changing household structure, enumeration and coverage could prove useful in efforts to maintain quality and coverage while reducing costs in the 2020 Census.

---

<sup>8</sup> Lamas, E. (2009), "Residence Rule and Residence Situations for the 2010 Census." Memorandum from Enrique Lamas, Chief, Population Division, to Frank Vitrano, Chief, Decennial Management Division, April 24, 2009. Suitland, MD: U.S. Census Bureau.

<sup>9</sup> Humes, Jones, and Ramirez, (2011), Overview of Race and Hispanic Origin: 2010. 2010 Census Briefs, C2010BR-02. Washington, DC: U.S. Census Bureau

<sup>10</sup> Schwede, L. (2007), A New Focus: Studying Linkages among Household Structure, Race/Ethnicity, and Geographical Levels, with Implications for Census Coverage." SRD Report RSM 2007/38. Suitland, MD: U.S. Census Bureau. Paper available at: <http://www.census.gov/srd/papers/pdf/rsm2007-38.pdf>.

<sup>11</sup> Elliott, Dye, and Young (2011), Variation in the Formation of Complex Family Households during the Recession. SEHSD Working Paper Number 2011-32. Washington, DC: U.S. Census Bureau.

<sup>12</sup> Parker, K. (2012), "The Boomerang Generation: Feeling Okay about Living with Mom and Dad." Pew Research Center Report. Released March 15, 2012.

<sup>13</sup> Mykyta, L. (2012), "Economic Downturns and the Failure to Launch: The Living Arrangements of Young Adults in the U.S. 1995-2011." SEHSD Working Paper 2012-24. Washington, DC: U.S. Census Bureau.

# 1. Introduction

## 1.1 Scope

The mission of the decennial census is to count everyone living in the United States once, only once, and in the right place. Accurate counts are important because census results are used over the following decade for allocating seats in the House of Representatives, state redistricting and distributing around \$400 billion in federal funds annually. Fulfilling this mission is daunting as the country grows more diverse and complex. The Census Bureau invests years in planning, research, and testing, and communications and hires more than 700,000 enumerators to collect data from persons who have not mailed their census forms. The decennial census is arguably the largest peacetime mobilization in the country.

Despite best efforts to count everyone, the Census Bureau's own record of research shows persistent differential undercounts of some minority populations across decennial censuses, such as Hispanics and African Americans. Statistical studies have quantified the types of errors and the characteristics of persons and households affected by them. Prior ethnographic studies conducted during census data collections, but separate from those operations, have identified a range of factors affecting coverage and have illuminated how and why they may affect enumeration in some populations. However, previous ethnographic studies have not included systematic observations of live interviews and respondents in an actual decennial environment.

The present evaluation was conducted to address this gap, in addition to addressing the overall need to increase the understanding of the types and sources of these persistent miscounts in the census. In this report, the authors present findings from comparative ethnographic and partial validation research. The goal of this evaluation was to explore types and sources of possible census coverage error and to identify the characteristics of households and of persons affected by them through observation of live 2010 Census interviews. We aimed to identify similarities and differences across race/ethnic groups to contribute to the discussion of why differential counts persist and to suggest improvements and new research for the next census.

To the authors' knowledge, this is the first decennial census in which researchers have been able to conduct systematic ethnographic observations of enumeration methods and coverage by observing interviewers conduct live interviews with respondents within actual data collection operations in the decennial census environment. This enables ethnographers to document factors within the interview that may have affected coverage, as well as debrief respondents in real time.

In this report, we follow the goals of the evaluation to identify types and sources of inconsistencies across records and possible coverage error, the characteristics of persons and households that may be affected by them, and how they may be related to race/ethnic groups that are chronically undercounted and miscounted. By having ethnographers observe live Census Coverage Measurement (CCM) interviews and debrief respondents immediately following the interview and validating these observations with a records check *post hoc*, we can examine types and sources of record inconsistencies and possible coverage errors and characteristics of persons and households affected by them by race and ethnic group. The records check study and data

from the 2010 Census and CCM provide quantitative evidence of the groups that are miscounted and suggestions of why. The qualitative insights from the ethnographers who actually observed the enumeration of these groups shed light on why some groups may be consistently undercounted and miscounted. This report starts with results on inconsistencies and possible coverage errors from the records check. It then summarizes crosscutting themes and profiles from each of the sites from the qualitative reports and interview summaries on reasons for inconsistencies and possible coverage errors.

This is a multi-stage, mixed methods evaluation. The first stage consisted of controlled comparisons of small-scale qualitative field observation studies across nine race/ethnic groups in which ethnographers accompanied interviewers to observe and audiotape (where permitted) live interviews and conduct immediate debriefings with respondents on coverage-related questions. The ethnographer observations were an independent assessment of where persons likely should be counted, reasons for any apparent miscounts, and other important enumeration issues. The second stage of this evaluation was a partial validation check of the census by identifying inconsistencies in where persons should be counted across data sources in a *post hoc* records check.

These stages were conducted on three personal-visit operations in 2010. Two of these operations were 2010 Census operations: Update Enumerate (UE) and Nonresponse Followup (NRFU). The third operation was the independent CCM Person Interview (PI) operation conducted some months later. The intent of observing both 2010 Census and later CCM operations in the same site was to facilitate the identification of wider factors in the sites over the time span of data collections that may have affected personal visit data collections. The initial design specified that the same ethnographer would observe both the census and later CCM operation in the site to document changes in the site and in different sets of respondents over time. However, that design was changed to have different ethnographers in the census and CCM observations in each site to avoid contaminating the independence of the CCM from the census, which was a core requirement in the CCM dual systems estimation methodology. The evaluation methods, procedures, and research questions were the same in each of the three operations.

These observation studies were conducted in sites specially selected by the evaluation authors to target primarily, but not exclusively, certain race/ethnic groups: African Americans in Chicago; Alaska Natives in Kodiak Borough, Alaska; American Indians on two Southwest reservations; Asians (primarily Chinese) in the San Francisco Bay area (including Chinatown during the NRFU operation); Native Hawaiian and Other Pacific Islanders on the Big Island of Hawaii; non-Hispanic Whites in the Kansas City, Missouri area in Jackson County; Hispanics in Dallas/Fort Worth, Texas; Middle Easterners in Wayne County Michigan (during CCM only); and a final generalized site in Broward, County, Florida.

There were two sets of ethnographers per site. The first set observed census operations in either the UE operation in March and April or the NRFU operation in mid- to late-May, 2010 during the height of the operations. Ethnographers for the NRFU/UE operations observed 363 interviews. The second set of ethnographers observed CCM PI operation in August, 2010 within specific pre-selected CCM sample clusters in these sites in August, 2010. Ethnographers for the CCM operations observed 318 interviews.

The second stage of this evaluation was a partial validation records check. The purpose of the records check was to partially validate census counts by triangulating across multiple data sources to identify inconsistencies in who should be counted at the ethnographer-observed housing units and to identify types and sources of possible coverage error. A records check inconsistency is defined as when the location of where a person should be counted differs between two or more data sources. Census Bureau staff matched and compared the rosters from 1) the standard interviews; 2) the ethnographers' assessments of where each person should be counted; 3) the final localized post-processed 2010 Census unedited dataset created for this evaluation, and, for the CCM observations, 4) the final localized post-processed 2010 CCM dataset formerly matched to census unedited records for CCM analysis, localized around the research sites for this evaluation. Taking into account data from all of these sources, the research team coders made their own assessment of where each person should be counted, which constituted the fifth data source.

The findings in this evaluation from observations of live personal visit interviews have direct relevance to 2020 Census working groups on questionnaire design, reducing and improving followup operations, and on developing adaptive design models. The Census Bureau faces greater organizational challenges for the 2020 Census in maintaining or improving census coverage and quality in the face of declining federal budgets that require rethinking and redesigning data collection methods. It is hoped that findings in this report contribute to those goals.

## **1.2 Intended Audience**

This report is intended for use by the sponsors and reviewers of this evaluation as well as program managers and staffs responsible for planning the 2020 Census, particularly the working groups for Reducing and Improving Followup, Questionnaire Design, and Adaptive Design. This report may be of use to those evaluating the 2010 CCM and planning the 2020 Census post-enumeration survey, and to Census Bureau Advisory Groups. It may be of interest to academics, researchers, policymakers, students, and members of the lay public.

## **2. Background**

### **2.1 History of Differential Coverage and Ethnographic Studies**

Persistent differential counting of minority populations is a challenging issue the Census Bureau faces in trying to fulfill its census mission of counting everyone once, only once, and in the right place. At the time we were formulating this evaluation's study plan to explore issues related to differential undercounts, we knew, for example, that African Americans had an estimated undercount of 4.57 percent in 1990 (Hogan, 1993) and 1.84 percent in Census 2000, while the comparable figures for Hispanics were 4.99 percent in the 1990 Census and 0.71 percent in Census 2000 (Hogan, 1993; May 22, 2012 Census Coverage news conference). American Indians had an estimated undercount rate of 12.22 percent in 1990, though after the Accuracy

and Coverage Evaluation II revisions, had a slight overcount that was not statistically significant from zero in Census 2000 (Mule 2003, U.S. Census Bureau 2003: Table 1; National Research Council, 2004). The comparable figures from these sources for non-Hispanic Whites showed an estimated undercount in 1990 of 0.68 percent but an overcount of 1.13 percent in Census 2000.

In May, 2012, after our data collection was completed, the final 2010 CCM results were released. Table 1 shows that similar patterns of differential undercounting persisted into the 2010 Census in which we had conducted this evaluation to explore factors associated with this phenomenon. Table 1 displays CCM results for the 1990 Census, Census 2000, and the 2010 Census (U.S. Census Bureau, 2012a and b). This shows significant Black undercount rates from 1990 through 2010. In contrast, non-Hispanic Whites had an undercount in 1990 but a significant overcount in 2010. Significant undercounts for Hispanics and for American Indians on reservations were recorded in the 1990 Census and in the 2010 Census.

Table 1: Percentage Net Undercount by Race and Hispanic Origin (National Census Coverage Measurement Estimates)

| <b>Race/Ethnic Group</b>               | <b>1990 Census</b> | <b>Census 2000</b> | <b>2010 Census</b> |
|--|--------------------|--------------------|--------------------|
| Non-Hispanic White                     | 0.68*              | -1.13              | -0.84*             |
| Non-Hispanic Black                     | 4.57*              | 1.84*              | 2.07*              |
| Non-Hispanic Asian                     | 2.36*              | -0.75              | 0.08               |
| American Indian On Reservation         | 12.22*             | -0.88              | 4.88*              |
| American Indian Off Reservation        | 0.68*              | 0.62               | -1.95              |
| Native Hawaiian/Other Pacific Islander | 2.36*              | 2.12               | 1.34               |
| Hispanic Origin                        | 4.99*              | 0.71               | 1.54*              |

Source: Excerpt from U.S. Census Bureau, 2012b table. Undercounts are shown as positive numbers, overcounts with minus signs. \* = Statistically significantly different from zero.

What factors have been associated with differential counting in past research? Past research has identified factors associated with undercounting, such as: residence rule issues (e.g., National Research Council, 2006; Martin, 2007; Gerber, 1994); mistrust of government and confidentiality concerns (Gerber, 2003); large households, irregular housing, and complex household structure (e.g., de la Puente, 1993; Schwede, 2003; Schwede, 2004; Schwede 2006; and Schwede, Blumberg, and Chan, 2006) and relationships as distant relatives or nonrelatives (Olson and Sands, 2012; Bell, 2004); mobility (Hunter, de la Puente, and Salo, 2003); renter status (Hogan, 1993; Ellis, 1994); Indian reservations and personal visit enumeration (Ellis, 1994; de la Puente, 1993) and colonias (de la Puente and Stemper, 2003). Other factors include motivational barriers, including deliberate concealment (Valentine and Valentine, 1971); distrust of the government and confidentiality pledges (Martin, Brownrigg, and Fay, 1990); language barriers and difficult access to housing units (de la Puente, 1993) and differences in respondent’s conceptual definitions of “household” and “usual residence” that can result in enumeration errors (Schwede, Blumberg and Chan, 2006; Bates and Gerber, 1998). (See also Appendix A).

This evaluation grew out of specific requests made by two Census Bureau Regional Office Directors when one of the authors asked each of them in 2005 what research questions they would like researchers to address in future qualitative studies. They asked for research to identify factors related to enumeration methods and coverage with special interest in doubled-up households. Based on their expert knowledge and first-hand experience in directing enumeration, the former Denver Regional Office Director called for research on why American Indians in public housing and on reservations underreport persons and additional families in their households, even when American Indian enumerators are used; her successor asserts that the issue remains critical. The Charlotte Regional Office Director called for research to meet challenges of correctly enumerating large and doubled-up Hispanic households which were increasing rapidly in his region and in others. The methods used in this evaluation also grew out of an earlier study that combined qualitative observations of live interviews during the 2006 Census Test on the Cheyenne River Sioux Indian reservation with a records check observation that identified types and sources of coverage error (Schwede, 2008a). That observation study on the Indian reservation was part of a larger observation study that also included observations in the Texas site during the 2006 Census Test (Nichols et al., 2007) which were based on the systematic observation of live interviews (Nichols and Childs, 2007; Nichols and Childs, 2009).

At the request of the 2010 Census Program for Evaluations and Experiments (CPEX) Executive Steering Committee in its January 2008 meeting, the scope of this study was expanded from just Hispanics and American Indians to cover all of the major race/ethnic groups.

## **2.2 Organization of this Report**

The methodology and limitations discussions (Section 3 and Section 4) below cover the census operations and then the CCM operation. Because the overall census and CCM data collection operations differ fundamentally in purpose and methods and what can be analyzed from them, Section 5.1 is devoted solely to the census operations and Section 5.2 to the CCM operation. From Section 6 onward, both are discussed.

## **3. Methodology**

The observation studies were conducted during three personal-visit data collection operations. In Section 3.1, the two 2010 Census operations and the questionnaire and other materials used will be discussed. In Section 3.2, the independent CCM and its purpose, instrument and other materials will be described.

### **3.1 Description of the 2010 Census and Census Coverage Measurement Operations**

#### **3.1.1 Update Enumerate Operation**

The UE operation is conducted in selected remote areas where city-style addresses are lacking and many households have only a post office box number that does not allow geocoding of their actual location. There is no mailout of census forms; enumerators canvass each census block in assignment areas, verifying and updating the addresses on the address lists and their locations on

the block maps. The enumerators also conduct interviews with persons in housing units in a personal visit or telephone interview (U.S. Census Bureau 2009b).<sup>1</sup>

The paper-and-pencil-based Enumerator Questionnaire (EQ) is used in the UE as well as in the NRFU operations. This form includes requests the respondent's name and phone number and 15 additional questions. It contains the questions on the mailout census form—tenure and respondent-provided population count of those living or staying in the housing unit on April 1, 2010—and a list of those persons and their demographic characteristics: name; relationship to Person 1; sex; age and birthdate; Hispanic/Latino/Spanish origin; race. It includes question 7 to flag possible erroneous enumerations: whether each person sometimes lived or stayed somewhere else in seven types of places/situations (in college housing; in the military; at a seasonal or second residence; for child custody; in jail or prison; in a nursing home; for another reason; or no). Like the mailout form, the EQ included a household-level omission question, “We do not want to miss any people who might have been staying here on April 1, 2010. Were there any additional people that you didn't mention, for example: Babies? Foster children? Any other relatives? Roommates? Any other nonrelatives? How about anyone else staying here on April 1, 2010 who had no permanent place to live?” If the respondent answered yes to any of these probes, he/she was asked to provide up to two persons' names.

The EQ includes additional questions: (S1) address confirmation, (S2) whether respondent or someone in the household lived or stayed in the observed housing unit on April 1, 2010 (used to determine if someone currently living in the housing unit can respond for the housing unit on April 1, 2010); (S3) whether the housing unit is usually occupied or is a vacation or seasonal residence; and (S4) whether the housing unit was occupied or vacant on April 1, 2010, as well as question H3 for entering an address if the form did not have an affixed address label. The EQ also included space for notes, for respondent type, interview summary, and certification.

A two-sided Information Sheet was used as a handout in both census operations. It included general information on confidentiality and the legal requirement to respond and four short lists keyed to questions in the form. List A gives selected residence rule instructions on whom to include and not include in Question S5 for the respondent count of those living or staying in the housing unit on Census Day. List B shows response categories for relationships, List C shows types of Hispanic, Latino or Spanish origins, and List D lists race categories, including “Some Other Race.” Update Enumerate was conducted from mid-March to mid-May, 2010.

### **3.1.2 Nonresponse Followup Operation**

The NRFU operation was designed to enumerate persons in housing units included in the Mailout/Mailback operation for which completed census forms were not checked in by the mailout cutoff date on April 19, 2010 (U.S. Census Bureau, 2009a).<sup>2</sup> About 26 percent of all

---

<sup>1</sup> The 2010 Census Update Enumerate Operations Assessment Report (Fallica, Heimel, Jackson, and Zhang 2012) provides more information on the UE operations.

<sup>2</sup> The 2010 Census Nonresponse Followup Operations Assessment Report (Walker, Winder, Jackson, and Heimel, 2012) provides more information on the NRFU operations.

housing units nationwide were included in the NRFU universe, with variation across localities. In this massive operation, more than 700,000 temporary Census Bureau enumerators knocked on more than 47,000,000 doors from May to July, 2010. Enumerators conducted personal visit or telephone interviews with respondents, using the Enumerator Questionnaire described above.

### **3.1.3 Census Coverage Measurement Person Interview Operation**

The CCM was an independent post-enumeration survey designed to estimate net census coverage as a whole and for certain populations, including race and Hispanic populations. This survey was independent of the census, with its own listing, data collection operations, and processing. Census Coverage Measurement housing unit and person data were matched to housing units and persons in the final 2010 Census, and dual systems estimation was used to estimate coverage. The CCM sample in the 2010 Census included about 187,000 housing units around the country, less than half of one percent of all housing units.

The CCM Person Interview (PI) operation, conducted from mid-August to October, 2010, was the first to collect information on persons in households. Interviewers conducted personal-visit or telephone interviews with respondents, using a computer-assisted person interview (CAPI) instrument.

The CCM instrument collected an inclusive roster of persons in housing units with special probes prompting respondents to identify persons who might have been left off the initial roster of those identified as living or staying in the housing unit. Probes included: anyone with another place to live who stays here often; anyone staying here until they find a place to live; babies, foster children or other children not mentioned yet; any relatives or unrelated people who live or stay here; and anyone living or staying here in March or April, 2010 who is no longer living here. The instrument also included questions to identify types of persons on the roster who perhaps should not be counted in the sample households according to the census residence rule: college students living away; persons in the military; those living or staying somewhere else for a job; those with a seasonal home; those in a group quarters around April 1, 2010; and anyone who stayed at another place often in the last year. Other questions asked for the length of stays elsewhere and patterns and durations of movements to decide where each person should be counted around Census Day.

## **3.2 Questions to be Answered**

The purpose of this study was to explore types and sources of possible census coverage error and identify the characteristics of households and of persons affected by them through observation of live 2010 Census interviews. By having ethnographers observe live interviews and conduct immediate debriefings with respondents on coverage ambiguities and omissions and then attempting to validate the census results in a *post hoc* records check, the authors can examine types and sources of possible coverage errors and characteristics of households and persons affected by them by race and Hispanic origin across all sites and also within sites. The authors aimed to identify similarities and differences across race/ethnic groups to contribute to the discussion of why differential counts persist and to suggest improvements and new research for the next census.



The research questions that we address in this report<sup>3</sup> include:

**3.2.1 What are the numbers, types and reasons for record inconsistencies<sup>4</sup> and possible coverage errors identified in our records check study?**

**3.2.2 What are the characteristics of persons and households with record inconsistencies and possible coverage errors?**

**3.2.3 What reasons for record inconsistencies and possible coverage errors are found in themes that cut across the ethnographers' reports?**

**3.2.4 How are record inconsistencies and possible coverage errors distributed across the race/ethnic groups and sites?**

**3.2.5 What characteristics specific to each ethnographic site (documented in ethnographic profiles) contribute to the identified record inconsistencies and possible coverage errors?**

### **3.3 Site Selection**

The objective was to select nine research sites targeted to specific race/ethnic groups in which one set of ethnographers would conduct census field observations and a different set would conduct later CCM field observations. The authors also wanted to select areas that were moderately hard-to-count to increase the likelihood that the researchers would observe possible coverage errors. Lastly, the authors wanted to select sites that were spread out around the country. To meet these objectives, the team used several data sources to select sites.

The authors began by using custom versions of the CCM sampling frame in early 2010 that identified CCM cluster locations and the numbers of households in them. These sampling lists showed just those sites that in Census 2000 had at least 15 percent of households of the targeted race/ethnic groups. Using these sampling frame lists by race/ethnicity, the authors short-listed sites with at least 100 households in the CCM sample that were not too dispersed, and at least 50 households with respondents in the targeted race/ethnic group for each site.

The authors used two sources to whittle down the list of possible sites to those tracts at higher risk of coverage error. The first was the Census Bureau's Planning Database hard-to-count scale

---

<sup>3</sup> The original study plan for this evaluation had six research questions on enumeration methods and coverage (Schwede 2011a, shown in Attachment B). During the preparation of this report, the decision was made that restructuring, combining, and reordering the questions to focus on types and sources of possible coverage error as the foremost theme would be a more effective and parsimonious way to present the findings. The Census Bureau's Decennial Management Division sponsors approved these changes.

<sup>4</sup> A records check inconsistency is defined as when the location of where a person should be counted differs between two or more data sources. Records check inconsistencies do not necessarily imply potential coverage errors. Census staff matched and compared the rosters from four data sources: 1) the standard interviews; 2) the ethnographers' assessments of where each person should be counted; 3) the final post-processed local 2010 Census unedited dataset and, for the CCM observations, 4) the final post-processed 2010 Census local CCM dataset matched to census records.

scores across 10 strata (Robinson, Bruce, Mills and Love, 2008). This database provided a hard-to-count score for each census tract, based on variables correlated with nonresponse in Census 2000. The team chose sites with moderate to high Planning Database hard-to-count scores.

The second source was the set of special segmentation clusters developed from Planning Database scores in the 2010 Integrated Communications Program to map the U.S for communications and advertising purposes (Bates and Mulry, 2011). Three of the eight clusters were of special interest: the “Ethnic Enclave,” “Economically Disadvantaged” and “All around Average” clusters with primarily renters. The team chose possible sites for the American Indian, Alaska Native, Native Hawaiian, Asian, and Hispanic observations from the Ethnic Enclaves segment, the African American site and the non-Hispanic White site from the “Economically Disadvantaged” segment, and the Generalized site from the “All Around Average” segment.

At the authors’ request, custom runs of tracts in these clusters that had at least 15 percent of each of the target race/ethnic groups and at least moderate tract-level hard-to-count scores, with the actual score for each tract, were created. The team matched the short-list of CCM sample sites for each race/group to the custom segment runs and cut the list of potential sites.

To make the final selection of American Indian reservations, the team consulted with administrators and the American Indian Partnership specialist in the Denver Regional Census Center. They also consulted with three American Indians in headquarters who worked on American Indian enumeration issues.

As noted above, the team’s design called for the inclusion of a generalized site not targeted to any specific group as a quasi-control site against which to compare the targeted sites. The team decided that a site that back in Census 2000 roughly mirrored the estimated national race/ethnic distribution in early 2010 could be considered roughly representative. A special tabulation was prepared to include only those tracts in the “All around Average” segment that had roughly 70 to 75 percent non-Hispanic Whites, roughly 10 to 15 percent African American and roughly 12 to 18 percent Hispanic. That list of tracts was then compared to the CCM sampling frame and one site was chosen. The final general locations of the race/ethnic sites are shown in Table 2.

Table 2: Locations of Race/Ethnic Research Sites

| <b>Race/Ethnic Group</b>               | <b>Location</b>                               |
|--|---|
| African American                       | Chicago, Illinois                             |
| Alaska Native                          | Kodiak Island Borough, Alaska                 |
| American Indian                        | Two Southwest Indian Reservations*            |
| Asian (mostly Chinese)                 | San Francisco Bay Area, California            |
| Middle Eastern**                       | Wayne County (Detroit and Dearborn), Michigan |
| Native Hawaiian/Other Pacific Islander | Big Island, Hawaii                            |
| Non-Hispanic White                     | Jackson County (Independence), Missouri       |
| Hispanic                               | Dallas/Fort Worth, Texas                      |
| Generalized site                       | Broward County, Florida                       |

Notes: \* = NRFU included just one reservation, CCM had two; \*\* = CCM only (see Section 3.3.)

Source: Comparative Ethnographic Studies: Observation Sites for CCM by Race/Ethnic Groups

### **3.4 Ethnographer Selection**

To implement this research design, the team needed two trained observers for each of the nine sites, for a total of 18 observers. Because staffing constraints at that time precluded using existing staff or hiring new ones for this project-specific work, the team chose to grant contracts to outside expert ethnographers who had prior research experience with the race/ethnic groups.

Recruitment flyers were used to identify potentially interested expert researchers with ethnographic skills, recent prior research experience in one or more of the target race/ethnic groups in the United States, and fluency in the target group languages when appropriate (for the Hispanic, Asian, and Middle Eastern sites). Flyers were distributed at professional association meetings and on targeted listserves of the American Anthropological Association and the American Sociological Association.

After receiving responses from more than 120 social scientists, the authors were able to match ethnographers to 17 of the race/ethnic sites across the census and CCM operations (Schwede and Terry 2012b). These ethnographers included two fluent Spanish speakers, a fluent Mandarin Chinese speaker, and one fluent Arabic speaker. Most were established researchers with Ph.Ds. in anthropology, sociology, or related fields and years of direct field research experience in the US with one or more of our target race and ethnic groups. Two were doctoral candidates who had already done research with their target groups and who were uniquely qualified to conduct a site evaluation, through a combination of language fluency and publications.

The team was unable to find a second Arabic speaker for the NRFU operation, thus that site was dropped, resulting in eight NRFU sites. A fluent Cantonese Chinese speaker could not be found for the CCM PI Asian site. Rather than lose the CCM PI Asian site, one of the co-authors conducted that observation study, despite her lack of knowledge of Chinese language and culture. This turned out unexpectedly well in the field because the sample was not concentrated in a Chinatown area; most respondents spoke some English (Schwede, 2011b).

### **3.5 Training**

The authors conducted three-day training sessions in 2010 at Census Bureau headquarters with the contract ethnographers in each operation: one for the ethnographer in the UE operation in March, sessions for the NRFU ethnographers in May, and sessions for the CCM PI ethnographers in August.

Each session included an overview of the operation to be observed, the standardized methods the researchers were to use, and copies of the questionnaire (for the 2010 Census) or summary of the automated instrument (for CCM PI), and relevant handouts and letters used by the interviewer in that operation. The team also gave the ethnographers operation-specific handouts to give to local field contacts and interviewers providing an overview of the evaluation and a set of bullets showing the respective roles and responsibilities of the observers (ethnographers) and the interviewers while conducting field interviews. Training included discussion of the principles of the census residence rule and situations document (Lamas, 2009), and emphasis on the

importance of ethnographers becoming familiar with the rules on how to determine where each person should be counted in the census. This would enable them to identify cues of possible coverage error while observing live interviews and then ask *ad hoc* debriefing questions at the end of the interview to determine where each person should be counted in the census.

In each training session, the authors played tape recordings of actual interviews with coverage error, identified the specific cues of possible coverage error and then played and discussed the actual debriefing that took place. In some training sessions the ethnographers role-played as an interviewer, observer, and respondent to help them become familiar with the interview situation and their roles in it vis-à-vis the interviewers and respondents.

### **3.6 Field Work**

The team developed standardized methods to conduct the field observation studies. One ethnographer was assigned to each of the sites in each operation according to previous experience with each racial or ethnic group. Each ethnographer aimed to go out with a different interviewer each day, both to minimize observer effects on interviewer behavior and also to get a sense of the range of interviewer behavior within sites. Each ethnographer was to observe 35 interviews, with as many audio-recorded as possible, or as many interviews as possible over nine days, whichever came first. In those sites targeted to a particular race/ethnic group, each ethnographer aimed to get from 50 percent to 75 percent of their interviews with respondents from that group. To do so, the ethnographer was authorized to ask their field liaison to send them out with enumerators going to areas with a higher probability of contacting respondents in that race/ethnic group. Prior to the interview, they requested the respondent's permission to tape, and then remained unobtrusive as the interviewer conducted the interview, watching and listening for cues of possible coverage error, and taking notes.

If the ethnographer detected any cues of possible coverage error, he or she was to conduct a short ad hoc debriefing with the respondent immediately following the interview to gather enough information to decide where each rostered person should be counted, according to the census residence rule. The ethnographer was also to follow up on any cues that a person who perhaps should be counted at that housing unit might have been omitted by the respondent or mentioned by the respondent but not rostered by the enumerator. In the debriefing, the researcher was to be conversational and say that he or she just wanted to ask a few questions about what the respondent said for clarification purposes, without implying that there was any error or omission on the part of either the interviewer or the respondent.

Later, the ethnographers transcribed their tapes verbatim and summarized untaped interviews as fully as possible. They prepared individual household interview summaries documenting if and where each person should be counted in the census, including apparent correct enumerations, possible omissions, erroneous enumerations, and other errors. They included other observations they had on the enumeration and classified household structure into the complex/noncomplex household typology provided to them (Schwede, 2008b). The ethnographers then analyzed their cases. They wrote site reports documenting their findings on enumeration methods and the types and sources of possible coverage error, as well as the characteristics of households and persons affected by possible coverage error. They also contextualized their findings by describing the

geographical, demographic, economic, social, and cultural features of the site and identifying wider factors that may have affected the enumeration.

During the 2010 Census operations, ethnographers observed a total of 363 interviews across eight sites. They audiotaped 247 of these interviews and provided detailed summaries for an additional 34. The distribution by site is shown in Section 5.1.

Later, during the CCM PI operation, the second set of ethnographers observed 318 interviews, with 252 of those audiotaped. This distribution will be discussed in more detail in Section 5.2.

### **3.7 Records Check**

After the operations were complete, a coding/matching team comprised of Census Bureau staff matched observed household and person data from multiple data sources to the final localized 2010 Census datasets and/or final localized 2010 CCM datasets specially created for this evaluation to do a partial validation study of the census in the observed housing units. The purposes were to triangulate across data sources to identify consistencies and inconsistencies as to who should be counted at the observed housing units and to attempt to identify possible sources of error, using information from all sources on a case-by-case basis. A records check inconsistency was defined as when the location of where a person should be counted differs between two or more data sources. Inconsistencies are not necessarily potential coverage errors in the census.

To determine if and where each person should be counted in the NRFU and UE samples, the coding/matching team compared rosters for 786 persons in 257 housing units that were occupied on April 1, 2010 and also had sufficient data from three data sources: 1) the roster collected in the observed standard EQ interview; 2) the roster of persons the ethnographer assessed should be counted there and elsewhere; and 3) the roster included in the final localized 2010 Census unedited file in the special local dataset created for this evaluation. From now on, this will be referred to as the “final localized 2010 Census unedited dataset” to distinguish it from the overall national 2010 Census Unedited File. For the NRFU and UE samples, analysis was at the level of general inconsistency between any two of the three data sources. Given all record sources, coders identified possible sources of inconsistencies across data sources and then recorded their best assessments as to where the person should be counted, adding the fourth data source to the analysis. Some cases were reviewed independently and sometimes recoded in team consultations. If information was insufficient, persons were coded as unresolved. Results indicated that 103 persons in 44 households had inconsistencies.

For those 953 persons in 288 households in the CCM PI operation, the coding/matching team compared rosters from four data sources: 1) the roster from the standard CCM PI interview; 2) the roster of persons the ethnographer assessed should be counted there and elsewhere; 3) the roster in the final 2010 Census unedited file in our special local dataset; and 4) the roster in the 2010 CCM Person Matching Review and Coding System (PerMaRCS) in the special local dataset developed for this evaluation, which may or may not have identified the same persons as the final 2010 Census unedited file roster. From here forward, this will be called the final localized 2010 CCM/Census matched dataset” to distinguish it from the overall national CCM dataset.

For the CCM sample, possible coverage errors were identified by comparing where the team determined the person should be counted – taking into account all data sources – with where the final 2010 Census unedited file in the special local dataset determined the person should be counted, adding the fifth data source to this analysis. Sixty persons with possible coverage errors were found in 41 observed CCM housing units, with 11 persons unresolved and nine with missing data.

#### **4. Limitations**

We note several limitations that apply to both the census and CCM operations. First, it is important to note this study was a qualitative study with purposive samples of race and ethnic sites around the U.S.. Convenience samples of respondents and housing units were also collected within each race and ethnic site with respondents who happened to be at home and willing to participate during the cold-call unannounced visits during the nine days the ethnographers observed interviews. The ethnographers were to remain passive and allow the interviewer to determine where and when to conduct interviews at which households. Thus, results from this evaluation are not generalizable beyond the race and ethnic group sites and respondents encountered in this evaluation.

In this study, caution is used when making statements about coverage error. All categories of the coverage error variable indicate *possible* coverage errors. It is possible that neither the team assessment nor the ethnographer assessment was the most informed on where a person should be counted. Some persons' records may include updates from followup interviews after the ethnographer-observed interviews, which may result in a more informed determination for where the person should be counted.

One limitation applies to the ethnographers' assessments. Ethnographers' assessments of correct enumeration and possible coverage error were limited to what they observed during the standard interview and what they asked during their debriefings. Ethnographers made their observations within the early weeks of data collection in each operation and would not know about any subsequent visits to their observed housing units later to follow up on missing or incomplete data or on anomalies. The team used their summaries and transcripts of what respondents actually said and did for added insights into possible coverage errors in observed households, but these transcripts and summaries are time-bound and do not reflect the results of any subsequent revisits to the households or Coverage Followup (CFU) interviews. Further, the final post-processed census results could have been affected by errors in checking in, transporting, imaging, or processing of forms.

Another factor is the limited geographical extent of the census and CCM final datasets that were used in the records check. They were custom-designed datasets designed to cover the blocks in which the observed interviews were located as well as two concentric rings of blocks surrounding those blocks, following standard CCM procedures before doing national searches. In this evaluation, coders/matchers searched for people within these final localized datasets surrounding the observed housing units, but did not match nationally. It is thus possible that

some number of the persons who were classified as possibly omitted from the 2010 Census in this report may have been included in the 2010 Census in some other area outside the geographical scope of our records datasets. It is also possible that some persons were duplicated in other parts of the country. Those situations are beyond the scope of this records check.

#### **4.1 2010 Census Operations**

It is important to note a special limitation in analyzing possible coverage error for the Census operations that collected just the data on the Enumerator Questionnaire and during ethnographer debriefings and the comparison to the final post-processed dataset. Each of the three data sources has limitations in terms of what they can say about the “truth” of correct enumeration and coverage error for the observed housing units during census operations in this evaluation.

Due to limitations in what was collected during the standard census interview with the Enumerator Questionnaire, definitive statements cannot be made on whether persons in the observed 2010 Census operations were actually counted in the right place or if they have actual coverage error. This questionnaire has two questions to identify persons who might be at risk of coverage error: either of being erroneously enumerated in the observed housing unit when they should be counted elsewhere (Question 7) or omitted from the census, or of being omitted from the unit when they were there on Census Day and should be counted there (H1). The limitation is that neither of these questions was designed to establish whether the persons did have coverage error; they were placed on the questionnaires to flag cases during processing that should be considered for inclusion in the subsequent 2010 Census CFU operation. The CFU telephone interview was a much more in-depth questionnaire designed to make final decisions on where each person should be counted. Those changes were already incorporated into the final localized 2010 Census unedited dataset and may sometimes have been sources of inconsistencies with the ethnographers’ assessments of where each person should be counted.

#### **4.2 Census Coverage Measurement**

Specifically for CCM, the authors were not able to hire a Cantonese-speaking ethnographer for the Asian site in the CCM operation. In response, one author, an anthropologist and survey methodologist, acted as a replacement and conducted the observation study in the Asian site so it would not be lost in this evaluation.

Another limitation pertains to the CCM instrument. Although the CCM automated survey instrument is specifically designed to identify coverage errors and determine if and where each person should be counted in the census, there may be some living situations that are not fully or accurately recorded by the instrument that might better be picked up through the ethnographer debriefings with the respondent.

One additional note is that the CCM sometimes picked up persons who were not identified in the 2010 Census and vice versa. We can document the number of those cases, but in some of them, we have no data on why persons may have been omitted.

Despite these limitations, the authors were able to conduct both the quantitative records check and also get very rich ethnographic information on most of our observed interviews, to present mixed-methods results on the types and sources of coverage error, who is affected, and how these differ across groups and sites. The authors present recommendations at the end for improving coverage during the research and development phase in preparation for the 2020 Census.

## **5. Results**

### **5.1 The 2010 Census Component**

This section presents the results from the 2010 Census NRFU/UE component of this evaluation. The section starts with a brief overview of the field observation phase in which ethnographers observed live interviews and debriefed respondents immediately after the interview. It documents the numbers of ethnographers and enumerators accompanied, as well as the numbers of interviews observed and audiotaped overall and with the target race/ethnic group, and shows variation across the race/ethnic sites. The focus then turns to the quantitative consistency analysis of pooled data across sites from separate data sources. By validating the final localized 2010 Census unedited dataset records with field observations and debriefings, we can examine types and sources of inconsistencies among records and characteristics of persons and households affected by it who may be at higher risk of coverage error. The focus then widens to identify cross-cutting themes from the insights in the ethnographers' site reports as well as profiles of factors in each race/ethnic site that shed light on why some groups may be consistently miscounted.

During the 2010 Census NRFU/UE operations, eight ethnographers accompanied 69 enumerators and observed a total of 363 live 2010 Census interviews across eight sites (the Middle Eastern site was not included in the Census observations), as shown in Table 3. Seven of the ethnographers accompanied 64 enumerators and observed 325 interviews in the second to fourth weeks of the Nonresponse Followup operation in May, 2010. The remaining ethnographer accompanied five interviewers and observed 38 interviews on the American Indian reservation during the UE operation from late March to early April; he was observing interviews right around April 1, 2010 while the others were observing six to eight weeks later.



Table 3: Interviewers Accompanied and Interviews Observed by Race/Ethnic Site

| <b>Race/ethnic Group</b>               | <b>Location</b>                    | <b>Interviewers accompanied</b> | <b>Interviews observed</b> | <b>Number of interviews with target race/ethnic group for site</b> |
|--|------------------------------------|---------------------------------|----------------------------|--|
| African American                       | Chicago, Illinois                  | 9                               | 35                         | 32   |
| Alaska Native                          | Kodiak Island Borough, Alaska      | 4                               | 78                         | 17   |
| American Indian                        | Southwest Indian reservation       | 5                               | 38                         | 35   |
| Asian (mostly Chinese)                 | San Francisco Bay Area, California | 13                              | 62                         | 59   |
| Native Hawaiian/Other Pacific Islander | Big Island, Hawaii                 | 5                               | 37                         | 23   |
| Non-Hispanic White                     | Jackson County, Missouri           | 11                              | 38                         | 24   |
| Hispanic                               | Dallas/Fort Worth, Texas           | 11                              | 40                         | 23   |
| Generalized site                       | Broward County, Florida            | 11                              | 35                         | Not applicable   |
|  |                                    |                                 |                            |  |
| <b>TOTAL</b>                           |                                    | <b>69</b>                       | <b>363</b>                 |  |

Source: Comparative Ethnographic Studies Ethnographer Reports

During the 2010 Census NRFU/UE operations, eight ethnographers accompanied 69 enumerators and observed a total of 363 live 2010 Census interviews across eight sites (the Middle Eastern site was not included in the Census observations), as shown in Table 3. Seven of the ethnographers accompanied 64 enumerators and observed 325 interviews in the second to fourth weeks of the NRFU operation in May, 2010. The remaining ethnographer accompanied five interviewers and observed 38 interviews on the American Indian reservation during the UE operation from late March to early April; he was observing interviews right around April 1, 2010 while the others were observing six to eight weeks later.

Also shown in Table 3, the ethnographers ranged in the total number of interviews they observed, from 35 in the African American and Generalized sites to 78 in the Alaska Native site, where the target population of Alaska Natives was small and dispersed in the site and the ethnographer tried to increase her likelihood of interviewing them. In this regard too, ethnographers varied in the proportion of all interviews they observed that were in households

with their target group; they ranged from 17 (22 percent) of 78 in the Alaska Native site to 59 (95 percent) of 62 in the Asian, primarily Chinese site in San Francisco Chinatown.<sup>5</sup>

A total of 247 (68 percent) of these were audiotaped. The percentages of all interviews that were audiotaped in each site is as follows: African American site (n = 28, 80 percent); Alaska Native site (n = 51, 51 percent) American Indian site (n = 21, 55 percent); Asian, mostly Chinese site (n = 24, 39 percent); Native Hawaiian/Other Pacific Islander site (n = 32, 86 percent); Non-Hispanic White site (n = 38, 100 percent); Hispanic site (n = 40, 100 percent); and Generalized site (n = 13, 37 percent).

The 247 audiotaped interviews that were transcribed verbatim and summarized provide rich data for coding and matching and statistical analysis. In 34 of the remaining untaped interviews ethnographers used extensive notes taken during the interviews or shortly thereafter to provide substantial interview reconstructions, summaries, or both. There was enough information in these reconstructions or summaries to include these cases in the overall dataset without adding substantial numbers of missing values to most variables. Interviews for four occupied housing units and two housing units with no population on Census Day in the non-Hispanic White site were dropped when it was discovered that the clusters in which these housing units were located were not part of the final localized 2010 Census unedited dataset, as they did not have an equal probability of being matched.<sup>6</sup>

---

<sup>5</sup> Some, but not all, of the ethnographers discussed the numbers of refused interviews. The ethnographer in the Texas site reported 26 persons refused of 229 contacted. Some were soft refusals in which respondents were busy or the right person was not at home and the enumerator was asked to return later. In the Native Hawaiian site, three refused: one reluctant respondent refused after the request to tape and two refused for political reasons having to do with mistrust of the federal government and the Hawaiian sovereignty movement. The ethnographer thought these men had decided not to participate before their arrival. The Alaska Native site had five refusals of 96 contact attempts. In the Asian site, there were nine refusals at 115 addresses contacted, with 42 notices-of-visit forms left when no one answered the door, and 62 completed interviews. In the Generalized site, one man refused outright, while two non-Hispanic whites were resistant, but completed the interviews as did two Hispanic respondents who were initially negative about participating.

<sup>6</sup> Of the remaining untaped cases, 69 observed interviews accounting for 220 persons either had a minimal ethnographer text-based summary that did not provide a clear enough picture of what happened in the interview or was missing a substantial amount of data or both. Some of these cases included names of persons with or without limited additional information, while others included just numbers of possible unnamed persons, presumably from proxy respondents. Coding and attempting to match these cases with minimal identifying information to the final local 2010 Census unedited dataset and adding them into the overall analysis dataset would have required more resources and time than was available and would have introduced more missing data into the dataset. For these reasons, it was decided to leave these 69 cases out of the effective analysis dataset. Of these 69 housing units, 18 were from the American Indian reservation (70 persons with names), 27 were from the Alaska Native site (103 persons), with about half having names, and 47 were from the Asian, almost exclusively Chinese site (47 persons), with 72 percent having names.

This yielded 275 observed housing units and 822 persons that had enough data to be used in the records check analysis. Eighteen more of the visited housing units were dropped because they had no population on Census Day.<sup>7</sup>

The resulting NRFU/UE analysis sample for the records check consisted of 257 ethnographer-observed housing units that had been occupied on Census Day. Of these 52 percent were reported as rented, 20 percent as owned with a mortgage or loan, 14 percent as owned free and clear and 4 percent as occupied without payment of rent. The remaining housing units had missing or otherwise unusable data.

In addition, the median NRFU/UE household size was 3.0 persons. One- and two-person households were the most frequent at 27 percent and 22 percent, respectively, followed by households with three persons (17 percent), four persons (15 percent), and five or more persons (20 percent). Two housing units had missing data.

Table 4 shows the NRFU/UE race and Hispanic origin distribution. In the analysis sample, 786 persons were identified in the 257 occupied NRFU/UE housing units as of April 1, 2010. Of these, 183 (23 percent) were Hispanics of any race according to the final localized 2010 Census unedited dataset records. The race distribution of the largest non-Hispanic groups is as follows: Black (15 percent); White (15 percent) and Asian (13 percent). Many persons in the occupied housing units were reported as being of multiple races (12 percent), almost all of them in the Native Hawaiian site.

The proportions of rostered males and females were equal. One person was coded as having multiple responses for the sex question.

The age cohorts in the NRFU/UE sample, ordered from largest to smallest proportions, were as follows: 30-49 years old (20 percent) , 18-29 years old (17 percent), 50 and older (16 percent), 10-17 years old (12 percent), 5-9 years old (7 percent), and 0-4 years old (7 percent). The remaining 21 percent had missing data for age.

---

<sup>7</sup> The distribution of the 20 housing units with no population on Census day was as follows: vacant (14), whole households with a usual home elsewhere (2), households in which all persons were in-movers after Census Day (i.e., April 1, 2010) (3), and nonresidential business (1).

Table 4: Distribution of Race and Ethnicity of NRFU/UE Persons in Occupied Housing Units across Race/Ethnic Sites

| Race/Ethnic Group Category                    | Race/Ethnic Sites |           |            |           |           |            |            |            |            |
|---|-------------------|-----------|------------|-----------|-----------|------------|------------|------------|------------|
|   | AA                | AI        | AN         | AS        | GS        | HI         | NH         | WH         | Total      |
| Hispanic White                                | 0                 | 0         | 3          | 0         | 19        | 32         | 1          | 4          | <b>59</b>  |
| Non-Hispanic White                            | 1                 | 2         | 22         | 2         | 15        | 0          | 13         | 61         | <b>116</b> |
| Non-Hispanic Asian                            | 2                 | 0         | 46         | 53        | 0         | 0          | 2          | 0          | <b>103</b> |
| Hispanic American Indian/Alaska Native        | 0                 | 0         | 1          | 0         | 0         | 0          | 0          | 0          | <b>1</b>   |
| Non-Hispanic American Indian/Alaska Native    | 0                 | 59        | 28         | 0         | 0         | 2          | 0          | 1          | <b>90</b>  |
| Hispanic Native Hawaiian/Pacific Islander     | 0                 | 0         | 0          | 0         | 0         | 0          | 2          | 0          | <b>2</b>   |
| Non-Hispanic Native Hawaiian/Pacific Islander | 0                 | 2         | 0          | 0         | 0         | 0          | 16         | 1          | <b>19</b>  |
| Hispanic Black                                | 1                 | 0         | 0          | 0         | 0         | 0          | 0          | 0          | <b>1</b>   |
| Non-Hispanic Black                            | 56                | 0         | 0          | 0         | 2         | 34         | 0          | 29         | <b>121</b> |
| Hispanic Other                                | 1                 | 1         | 2          | 0         | 8         | 62         | 5          | 6          | <b>84</b>  |
| Non-Hispanic Other                            | 2                 | 0         | 1          | 0         | 7         | 0          | 2          | 0          | <b>13</b>  |
| Hispanic Multiple                             | 0                 | 0         | 2          | 0         | 0         | 2          | 10         | 0          | <b>14</b>  |
| Non-Hispanic Multiple                         | 5                 | 0         | 6          | 1         | 3         | 3          | 50         | 4          | <b>80</b>  |
| <b>Total</b>                                  | <b>68</b>         | <b>72</b> | <b>111</b> | <b>56</b> | <b>54</b> | <b>135</b> | <b>101</b> | <b>106</b> | <b>703</b> |

Source: Comparative Ethnographic Studies NRFU/UE Analysis Sample Dataset

Note: 83 persons had missing data for combined Hispanic origin and race. AA = African American site, AI = American Indian site, AN = Alaska Native site, AS = Asian site, GS site = Generalized Site, HI = Hispanic site, NH = Native Hawaiian and Other Pacific Islander site, WH = non-Hispanic White.

### 5.1.1 Records check inconsistencies: numbers, types, and reasons

What are the numbers, types, and reasons for records check inconsistencies identified in the records check stage of this study? In this section, the analysis focuses on a partial validation of the census by triangulating data on where people were or should have been counted across separate data sources and examining inconsistencies between data sources. These inconsistencies may or may not be possible coverage errors, due to limitations of the EQ form in determining where each person should be counted. As stated in the limitations section, the Enumerator Questionnaire used in the UE and NRFU operations was designed to collect basic census information comparable to that collected on the mailout census form. While it does include questions to identify some persons who *might* have possible coverage error—such as those who might live or stay in more than one place (Question 7) and are possible erroneous enumerations, and persons who might potentially be omitted (Question H1) when they should be included—the EQ questions are insufficient in and of themselves for determining where *all* rostered persons should be counted. They were intended as flags for determining during processing which cases should go forward into the more in-depth 2010 Census CFU. The CFU, similar but completely separate from the CCM, is designed to determine where persons should be counted in the census. Its results are used to improve census coverage directly.

Despite these limitations, it is enlightening to look at the inconsistencies across records to identify factors and situations that can cause trouble in enumeration.

#### 5.1.1.1 Numbers and Types of Record Inconsistencies

Issues that may be related to coverage in NRFU/UE were investigated by analyzing the degree of consistency from multiple data sources of where persons likely should have been counted. For each person in NRFU/UE housing units that were occupied on April 1, 2010, coder/matcher team members coded where each person likely should have been counted based on the perspective of three data sources. The three data sources used in the comparison for this analysis were the: (a) EQ standard interview transcript or summary, (b) ethnographer assessment and (c) the final localized 2010 Census unedited data imported into the PerMaRCS system for geographical areas surrounding the observed housing units. The coder/match team member assessment, based on all available data, became the fourth data source for comparison. For the EQ standard interview transcript source, a coder/matcher team member read the ethnographer's transcript for the case and determined solely from data from the transcript where the person likely should have been counted. Data for the ethnographer's assessment came from the coder/matcher coding the ethnographer's summary of the case on where persons likely should be counted. The final localized 2010 Census unedited dataset shows where persons were counted according to the final processed 2010 Census Unedited File (CUF) that was pulled from the overall national CUF for concentric rings of blocks around the observed housing units. The coder/matcher team member assessment is where the coder/matcher determined the person likely should have been counted after taking into account data from the NRFU/UE Standard Interview Transcript, ethnographer's assessment, and the final localized 2010 Census unedited dataset.

Two types of consistency analysis were conducted. In the first analysis, the first three data sources listed above were compared to one another in three binary yes/no match comparisons:

1. NRFU/UE Standard Interview to Ethnographer Assessment
2. NRFU/UE Standard Interview to the final localized 2010 Census unedited dataset
3. Ethnographer Assessment to final localized 2010 Census unedited dataset

In the second analysis, the coder/matcher team member assessment was included, which resulted in six binary yes/no match comparisons for each person:

1. NRFU/UE Standard Interview to Ethnographer Assessment
2. NRFU/UE Standard Interview to final localized 2010 Census unedited dataset
3. Ethnographer Assessment to final localized 2010 Census unedited dataset
4. Team Assessment to final localized 2010 Census unedited dataset
5. Ethnographer Assessment to Team Assessment
6. Team Assessment to 2010 Census NRFU/UE Standard Interview

A separate analysis was conducted with the coder/matcher team member assessment because team members considered all other data sources when making their assessment for this variable, which thus generally made the team assessment relatively more informed than the other data sources taken individually.

Overall, the most consistent binary match was between the ethnographer assessment and the EQ standard interview, which matched 96 percent of the time. The least consistent binary match was between the ethnographer's assessment and the final localized 2010 Census unedited dataset, which matched 88 percent of the time.

It can be seen from Table 5 that the consistency match rates are lowest when the binary comparison includes the final localized 2010 Census unedited dataset rosters, with the ethnographer/final localized 2010 Census unedited dataset match rate the lowest, around 88 percent. A major reason for this is that some persons were found to be either in the observed interview or in the final localized 2010 Census dataset, but not both—these are potential census coverage errors that need to be examined on a case-by-case basis; some may not be resolvable.

Of the 103 persons who were inconsistent, 30 persons who were either identified in the standard EQ interview or by the ethnographer during the debriefing were not found in the final localized 2010 Census unedited dataset available to the coder/matchers in PerMaRCS. There are several possible explanations for this: 1) some of these persons were missed in the local final localized 2010 Census unedited datasets but might have been included in the census elsewhere in the country; 2) some may have been missed altogether in the 2010 Census; and 3) some perhaps should not have been counted in the 2010 Census, though this option is lessened by the

ethnographers' direct observations of the interviews. These are cases of special interest and some will be discussed in the ethnographic profiles section.

Conversely, 15 persons were identified in the final localized 2010 Census unedited dataset who were not identified during the ethnographer-observed standard NRFU/UE interview or in the debriefing: eight of these were in the three housing units in the non-Hispanic White site (with five from one interview; three were from one housing unit in the Hispanic site); two were from one housing unit in the American Indian site and one was from one housing unit in the Alaska Native site. The only data on these cases came from the final localized 2010 Census unedited files; these persons were not identified in the ethnographic observations so the ethnographers would have no knowledge of these people. These inconsistencies are of interest in terms of coverage, and some of these cases are described in the profiles.

Table 5: Rank Order of Consistency between Data Sources for NRFU/UE Persons

| <b>Data Source Binary Match</b>  | <b>Match Percentage</b> |
|--|-------------------------|
| NRFU/UE Standard Interview to Ethnographer Assessment                      | 96                      |
| Team Assessment to NRFU/UE Standard Interview                              | 95                      |
| Ethnographer Assessment to Team Assessment                                 | 95                      |
| Team Assessment to final localized 2010 Census unedited dataset            | 89                      |
| NRFU/UE Standard Interview to final localized 2010 Census unedited dataset | 89                      |
| Ethnographer Assessment to final localized 2010 Census unedited dataset    | 88                      |

Note: Percentages are rounded to the nearest whole number.

Source: Comparative Ethnographic Studies NRFU/UE Analysis Sample Dataset

The first analysis included just the first three data sources, for a maximum of three possible matches. Overall, 675 persons (87 percent) had three of three matches, which indicated that all three data sources matched each other in terms of where each data source determined the person likely should have been counted. Of the remaining 100 persons with enough data, 91 had one match and nine had zero matches. At the housing unit level, 44 of 257 occupied housing units had at least one person with an inconsistency across data sources.

The second analysis consisted of matching all four data sources, including the coder/matcher team member assessment that was omitted from the first analysis. This additional data source resulted in a maximum of six possible matches.

In this analysis, the majority of persons (87 percent) had six matches, or had all four data sources match with each other. For the remaining persons, six had three matches (8 percent), two matches (3 percent), or one match (3 percent). Those with the lowest number of matches would likely be the ones with the higher likelihood of having coverage error.

#### 5.1.1.2 Reasons for Record Inconsistencies

Coder/matcher team members coded each person in NRFU/UE housing units that were occupied on April 1, 2010 for sources of inconsistencies across data sources. Sources of inconsistency are characteristics of persons, households, or the overall survey process or other factors that can lead to possible coverage errors in surveys and censuses. Up to three sources of inconsistency were coded for each person. Of 103 persons who had at least one inconsistency across data sources, team members coded 25 different sources of inconsistency for a total of 161 times.

Table 6 shows that the most frequent sources of inconsistency were unknowledgeable proxies (n = 21) and question rewording (n = 20). “Cycling” in this table refers to persons who move back and forth between places, such as children moving between the homes of their divorced or separated parents “Cycling” also includes tenuously attached persons who cycle among various housing units and/or other places, and college students who cycle between college housing and their parents’ homes.

All team sources of inconsistency were also collapsed into eight broad categories. Of these broader categories shown in Table 7, interviewer error accounted for 29 percent of the sources listed, mobile or tenuous living situation for 12 percent of them and respondent concealment or refusal for 9 percent for NRFU/UE persons with at least one inconsistency.



Table 6: Team Sources of Inconsistency for NRFU/UE Persons with at Least One Inconsistency across Data Sources

| Source of Inconsistency  | Frequency  | Percent    |
|--|------------|------------|
| Unknowledgeable proxy  | 21         | 13         |
| Question rewording   | 20         | 12         |
| Omission of question(s)  | 14         | 8          |
| Not included on Census form but person mentioned during NRFU interview   | 13         | 8          |
| Incomplete interview   | 11         | 7          |
| Interviewer error: what was recorded in the Census records does not match what was recorded in NRFU interview        | 9          | 7          |
| Multiple housing units on property   | 8          | 4          |
| Respondent confusion   | 8          | 4          |
| No official address for housing unit   | 7          | 3          |
| Partially inaudible tape recording   | 6          | 3          |
| Respondent annoyance with the interview  | 6          | 3          |
| Extended stay/work detail  | 5          | 3          |
| Respondent ambiguity   | 4          | 3          |
| Inconsistent living arrangements   | 4          | 3          |
| Respondent refusal to answer   | 4          | 3          |
| Interviewer making assumptions   | 4          | 3          |
| Cycling for college  | 4          | 3          |
| Cycling for custody  | 3          | 2          |
| Confusion due to language barrier  | 3          | 2          |
| Possible miscount due to outmover status   | 2          | 1          |
| Duplicate records due to multiple Census operations  | 1          | 1          |
| Possible miscount due to inmover status  | 1          | 1          |
| Enumerator did not fully read questions due to field conditions (evasive respondent, uncooperative respondent, etc.) | 1          | 1          |
| Military deployment  | 1          | 1          |
| Not included in NRFU standard interview, but person listed in Census records   | 1          | 1          |
|  |            |            |
| <b>Total</b>   | <b>161</b> | <b>100</b> |

Note: Percentages are rounded to the nearest whole number.

Source: Comparative Ethnographic Studies NRFU/UE Analysis Sample Dataset

Table 7: Collapsed Team Sources of Inconsistency for NRFU/UE Persons with at Least One Inconsistency across Data Source

| Source of Inconsistency               | Frequency  | Percent    |
|---------------------------------------|------------|------------|
| Interviewer error                     | 47         | 29         |
| Mobility or tenuousness               | 20         | 12         |
| Respondent concealment/refusal        | 15         | 9          |
| Address missed during the Census      | 15         | 9          |
| Not in the Census for unknown reasons | 13         | 8          |
| Respondent confusion                  | 8          | 5          |
| Language barrier                      | 3          | 3          |
| Other                                 | 40         | 25         |
| <b>Total</b>                          | <b>161</b> | <b>100</b> |

Note: Percentages are rounded to the nearest whole number.

Source: Comparative Ethnographic Studies NRFU/UE Analysis Sample Dataset

### 5.1.2 Characteristics of persons and households with record inconsistencies

#### 5.1.2.1 Persons

As stated previously, the NRFU/UE sample was not analyzed for possible coverage error, and instead NRFU/UE persons and households were analyzed for consistency among the four data sources and assessments: (a) EQ standard interview transcript, (b) ethnographer assessment (c), local final 2000 Census unedited data imported to PerMaRCS, and (d) coder/matcher team member assessment. Of 786 total persons in NRFU/UE occupied housing units, 103 persons had at least one inconsistency across data sources.

Regarding age, the 18-29 age cohort was the largest cohort to have at least one mismatch across data sources (n = 19), although the 0-4 age cohort had the greatest proportion of inconsistency among all age cohorts, with 11 of 47 (19 percent) of 0-4 year olds with an inconsistency (compared to 11 percent for the sample as a whole).

Overall, males (n = 47) and females (n = 49) had similar frequencies of persons with an inconsistency, while the remaining seven persons had missing data or reported multiple responses. When cross-tabulated by age, the 5-9 age cohort (five males and three females) and the 30-49 age cohort (three males and eight females) had the greatest proportional difference between males and females with an inconsistency. The 18-29 cohort was relatively balanced by sex, with 11 males and 8 females with an inconsistency.

Besides the reference person (n = 28), the most frequent relationships to Person 1 were biological sons and daughters (n = 19). Grandchildren (n = 7) and nieces or nephews (n = 5) to Person 1 also had inconsistencies. Besides the adopted children and parent-in-law with one consistent person and one inconsistent person each, the relationship types with the greatest proportions of

persons with inconsistencies within their own types were nieces/nephews (39 percent),<sup>8</sup> grandchildren (33 percent), and stepsons/daughters (25 percent), and children for whom it could not be determined whether they were biological, adopted, or stepchildren (23 percent).

#### 5.1.2.2 Housing units and households

As previously stated, 44 of 257 housing units had at least one person with an inconsistency across data sources and assessments. Regarding tenure, most housing units with at least one inconsistency were rented ( $n = 18$ ) or owned either free and clear or with a mortgage ( $n = 15$ ). Although rented housing units had the most inconsistencies of all tenure types in the overall sample, the proportion of owned housing units with and without an inconsistency (35 percent) was greater than the proportion of rented households with and without an inconsistency (13 percent).

The most-frequent household size for persons with at least one inconsistency were one-person households ( $n = 13$ ) and five or more person households ( $n = 13$ ). Proportionally among households of the same household size, five or more person households (26 percent), four-person households (19 percent), and one-person households (19 percent) were the household sizes with the greatest proportions of inconsistency.

Past qualitative research has shown that coverage errors are more likely to occur in complex than noncomplex households (e.g., de la Puente 1993; Schwede 2006, 2004 and 2003; Schwede, Blumberg, and Chan, 2006). For the purposes of this analysis, households are classified into the complex/noncomplex household typology (Schwede 2008b).

Noncomplex households generally consist of three types of households. Two of these three are “family households,” with two or more persons, at least one of whom is related to Person One by birth, marriage or adoption. The first of these two noncomplex family household types is the “Ozzie and Harriet”<sup>9</sup> type of nuclear family, strictly defined as a married husband and wife, and any of their own joint biological children who all live together in the same household. The second noncomplex family type is a stem family, which consists of a parent and his/her biological children with no one else. The final noncomplex type is a type of nonfamily household: one person living alone.

Complex households comprise all other combinations of two or more persons. They include persons other than, or in addition to, nuclear family relatives. Complex family households include blended families with stepchildren; grandparents with grandchildren, and other linearly extended households as well as laterally extended households with adult siblings, aunts, cousins,

---

<sup>8</sup> Nieces and nephews cannot be tabulated from 2010 Census relationship data because they fell into the general “other relative” category and could not be disaggregated. The ethnographers identified a wider range of relationships from the standard interviews and debriefings. The numbers of nieces/nephews, grandchildren, and stepchildren were low.

<sup>9</sup> “Ozzie and Harriet” was a popular television show in the 1960s that focused on the lives of a married couple and their biological sons and came to typify the nuclear family. This term was used in some of the past literature on families and is still sometimes used.

nieces, and so on. They may also include one or more nonrelatives, such as an unmarried partner or a roommate, or more than one family.

Lastly, a complex nonfamily household includes two or more persons, none of whom is related to the householder. Because of the way relationships are referenced only to Person One in the 2010 Census and in CCM, it is possible in some nonfamily and family households to have hidden family relationships that are not detected in the relationship question if, for example, neither Person Three nor Person Four is related to Person One but they are related to each other.

There is evidence from the 2010 Census, Census 2000, and the 1990 Census that complex households are associated with coverage error. Recently released 2010 CCM results indicated that while householders, spouses and “own children” (biological, step and adopted children) had overcount rates of 0.32 and adult children had high overcount rates of 2.91, other household members had quite high undercount rates of 3.53 percent (Olson and Sands, 2012). Those would include more distant relatives and nonrelatives (complex households). Similarly, several prior logistic regression studies from the 1990 Census Post-Enumeration Survey and the Census 2000 Accuracy and Coverage Evaluation showed that other relatives and nonrelatives were less likely to be included in the census than Person One (Bell, 2004).

Of the 257 NRFU/UE occupied housing units, 82 were complex households, and 162 households were noncomplex. The remaining 13 households were excluded because they had household types that could not be determined by the coder/matcher team members.

Of the 44 housing units that had a person with at least one inconsistency across data sources and assessments, 19 consisted of noncomplex households and 17 were complex households.<sup>10</sup> The remaining eight were excluded because they had missing data for household type.

Chi-square tests of independence were conducted to test the null hypothesis of no association between household type and inconsistency between NRFU data sources. The chi-square test found no association between household type and inconsistency between data sources ( $X^2 = 3.50$ ,  $DF = 1$ ,  $p = .0610$ ).

### **5.1.3 Distribution of record inconsistencies across race/ethnic groups and sites**

Of persons who had at least one mismatch between data sources ( $n = 103$ , 13 percent of all cases), 38 percent were non-Hispanic Blacks and 26 percent were non-Hispanic American Indian and Alaska Natives. Non-Hispanic Whites (13 percent) and non-Hispanic Asians (12 percent) also had inconsistencies, but very few consistencies occurred for persons outside of these four groups, including Hispanics of any race. See Table 8 for the distribution of race and ethnicity among persons with at least one inconsistency across data sources.

Chi-square tests of independence were also conducted that individually compared persons of different race groups with one another, to test the null hypothesis of no association between any

---

<sup>10</sup> “Housing unit” refers to the physical unit, while “household” refers to the group of persons who share one housing unit. The numbers of housing units and households are the same in census tabulations produced by the Census Bureau.

of the race/ethnic groups and an inconsistency between NRFU data sources. American Indian and Alaska Natives of any race ( $X^2 = 10.54$ ,  $DF = 1$ ,  $p = .0012$ ) and African Americans of any race ( $X^2 = 22.23$ ,  $DF = 1$ ,  $p < .0001$ ) were the race/ethnic groups associated with inconsistency between data sources. This is not to say that there is anything inherent in being African American, American Indian or Alaska Native, or any other race group that is a causal factor in inconsistencies. There are likely factors associated with the living situations in these sites, the enumeration process, and the sources of inconsistencies mentioned earlier, and other factors that affect inconsistencies.

Table 8: Race and Ethnicity of NRFU/UE Persons with at Least One Inconsistency across Data Sources

| <b>Race/Ethnic Group</b>                  | <b>Total Persons in Group</b> | <b>Frequency of Persons with an Inconsistency</b> | <b>Percent Within Group</b> |
|---|-------------------------------|---|-----------------------------|
| Black                                     | 123                           | 24  | 20*                         |
| American Indian/Alaska Native             | 91                            | 16  | 18*                         |
| Asian                                     | 103                           | 7   | 7                           |
| White                                     | 176                           | 9   | 5                           |
| Hispanics of Any Race                     | 183                           | 5   | 3                           |
| Multiple Races                            | 94                            | 4   | 4                           |
| Other                                     | 97                            | 1   | 1                           |
| Native Hawaiian or Other Pacific Islander | 21                            | 0   | 0                           |

Note:  $p < .05$ . Eighty-one persons were excluded from analysis because they had missing data for combined race and Hispanic origin. Percentages are rounded to the nearest whole number. Source: Comparative Ethnographic Studies NRFU/UE Analysis Sample Dataset

Regarding site differences, the African American site ( $n = 28$ ) had the most persons with at least one inconsistency, followed by the Non-Hispanic White site ( $n = 17$ ) and American Indian site 1 ( $n = 16$ ). See Table 9 for persons with at least one inconsistency between data sources across race/ethnic sites. Chi-square tests of independence were conducted that individually compared each site with the other sites test the null hypothesis of no association between any one race/ethnic site and an inconsistency between NRFU data sources. Results show that the African American site ( $X^2 = 36.48$ ,  $DF = 1$ ,  $p < .0001$ ) and the American Indian site 1 ( $X^2 = 4.16$ ,  $DF = 1$ ,  $p = .0414$ ) were associated with inconsistency between data sources.

When interpreting the number of inconsistencies by race and by site, it is important to note a possible limitation with the ethnographers' assessment. The ethnographers in the African American site and Non-Hispanic White site coded a substantial number of their cases as unresolved in their reports relative to other sites. This was due in large part to the facts that they each documented large numbers of interviews in which enumerators did major rewording of some to all of the critical coverage questions and their attempts to recontact the respondents by

telephone for partial reinterview with the questions worded correctly were not successful in more than half of the cases. Recognizing that stricter coding of persons as unresolved by these two ethnographers could potentially be a factor in raising the number of inconsistent cases in their sites vis-à-vis the other sites, the authors and a team member did a special review of the non-Hispanic White site and African American site interviews. They independently reviewed those transcripts and wrote up individual assessments of the ethnographer’s codes and the team codes, then discussed and reached consensus on whether any should be recoded. In this way, the team attempted to reduce the likelihood that differing ethnographer coding standards affected the distribution of inconsistencies by sites.

Table 9: NRFU/UE Persons with Inconsistency across Data Sources across Race/Ethnic Sites

| Race/Ethnic Site         | Frequency of Inconsistency | Percent Between Sites | Percent Within Site |
|--------------------------|----------------------------|-----------------------|---------------------|
| African American         | 28                         | 27                    | 35*                 |
| American Indian (site 1) | 16                         | 15                    | 21*                 |
| Non-Hispanic White       | 17                         | 17                    | 14                  |
| Asian                    | 9                          | 9                     | 15                  |
| Native Hawaiian          | 13                         | 13                    | 12                  |
| Alaska Native            | 11                         | 11                    | 9                   |
| Hispanic                 | 7                          | 6                     | 5                   |
| Generalized Site         | 2                          | 2                     | 4                   |
| <b>Total</b>             | <b>103</b>                 | <b>100</b>            |                     |

Note:  $p < .05$ . Percentages are rounded to the nearest whole number.

Source: Comparative Ethnographic Studies NRFU/UE Analysis Sample Dataset

Regarding the sex of persons with an inconsistency across the race/ethnic sites, the African American site had more females ( $n = 18$ ) with an inconsistency than males ( $n = 9$ ). Conversely, the non-Hispanic White site had more males ( $n = 11$ ) with an inconsistency than females ( $n = 6$ ). Among the race/ethnic sites, the Non-Hispanic White site and the Asian site had the greatest sex difference in persons with an inconsistency, each with a five-person difference in males as compared to females with an inconsistency.

Table 10 shows that among NRFU/UE households included in analysis, the American Indian reservation site 1 had the highest proportion of complex households (65 percent), followed by the Hispanic site (41 percent). The sites with the lowest proportions were the African American site with 18 percent, the Asian site with 20 percent and the Generalized site with 22 percent. Most of the cases in those sites were in small apartments or single-room occupancy, units which may have limited the size of the households.

Table 10: Census Day Household Type for NRFU/UE households by Race/Ethnic Sites

| Race/Ethnic Site         | Noncomplex Households |                     | Complex Households |                     | Total      |
|--------------------------|-----------------------|---------------------|--------------------|---------------------|------------|
|                          | Total                 | Percent Within Site | Total              | Percent Within Site |            |
| Alaska Native            | 27                    | 66                  | 14                 | 34                  | <b>41</b>  |
| Non-Hispanic White       | 22                    | 60                  | 14                 | 39                  | <b>36</b>  |
| Native Hawaiian          | 24                    | 69                  | 11                 | 31                  | <b>35</b>  |
| African-American         | 23                    | 82                  | 5                  | 18                  | <b>28</b>  |
| Hispanic                 | 22                    | 60                  | 15                 | 41                  | <b>37</b>  |
| Generalized              | 21                    | 78                  | 6                  | 22                  | <b>27</b>  |
| Asian                    | 16                    | 80                  | 4                  | 20                  | <b>20</b>  |
| American Indian (site 1) | 7                     | 35                  | 13                 | 65                  | <b>20</b>  |
| <b>Total</b>             | <b>162</b>            |                     | <b>82</b>          |                     | <b>244</b> |

Note: Thirteen households were excluded from analysis because their household type could not be determined.

Source: Comparative Ethnographic Studies NRFU/UE Analysis Sample Dataset

Of the 44 housing units with at least one inconsistency across data sources and assessments, 17 were complex households (See Table 11). In the Native Hawaiian site, four of the five with at least one inconsistency were complex. In the American Indian reservation (site 1), three of five households were complex, and in the generalized site, the only household with an inconsistency was complex. In the Asian site and the Hispanic site, households with any inconsistencies were split evenly between complex and noncomplex. In contrast, in the remaining sites, households with inconsistencies were more often noncomplex.

The Native Hawaiian site (n = 5) and the American Indian site (n = 3) had the highest proportions of complex households with at least one inconsistency across data sources and assessments. Table 11 shows the distributions across race/ethnic sites.

Table 11: Census Day Household Type for NRFU/UE Households with at Least One Inconsistency across Race/Ethnic Sites

| Race/Ethnic Site       | Noncomplex Households |                     | Complex Households |                     | Total     |
|------------------------|-----------------------|---------------------|--------------------|---------------------|-----------|
|                        | Total                 | Percent Within Site | Total              | Percent Within Site |           |
| African-American       | 6                     | 86                  | 1                  | 14                  | 7         |
| Non-Hispanic White     | 3                     | 60                  | 2                  | 40                  | 5         |
| Alaska Native          | 4                     | 57                  | 3                  | 43                  | 7         |
| American Indian Site 1 | 2                     | 40                  | 3                  | 60                  | 5         |
| Hispanic               | 2                     | 50                  | 2                  | 50                  | 4         |
| Asian                  | 1                     | 50                  | 1                  | 50                  | 2         |
| Native Hawaiian        | 1                     | 20                  | 4                  | 80                  | 5         |
| Generalized Site       | 0                     | 0                   | 1                  | 100                 | 1         |
| <b>Total</b>           | <b>19</b>             |                     | <b>17</b>          |                     | <b>36</b> |

Note: Eight households with a person with an inconsistency were excluded because they had missing data for household type. Percentages are rounded to the nearest whole number.

Source: Comparative Ethnographic Studies NRFU/UE Analysis Sample Dataset

Table 12 shows that the sites with the highest frequencies of team sources of inconsistency were the African American site (45) and the non-Hispanic White site (25). Those with the lowest frequency of team sources of inconsistency were the Alaska Native site and the Generalized site. As mentioned, for the African American site and the non-Hispanic White site, the ethnographers originally coded many of their persons as “unresolved” if they were not able to conduct followup interviews in situations where the interviewer did not properly ask all parts of the key coverage questions and these were not probed immediately after the interview. The independent review of these reduced these codes somewhat. Nonetheless, the frequencies of persons with an inconsistency should be interpreted with some caution in the African American site and the non-Hispanic White site.

Regarding team sources of inconsistency across data sources, the most-frequent sources were unknowledgeable proxies and question rewording. Of team sources, question rewording was coded most often at the African American site (n = 16). As noted earlier, a number of interviews in this site had major rewording or omission of some key coverage questions. See Table 13 for specific team sources of inconsistency across race/ethnic sites.



Table 12: Team Sources of Inconsistency for NRFU/UE Persons with at Least One Inconsistency across Data Sources

| Site                     | Persons With Inconsistency Within Each Site |         | Different Types of Sources Coded | Frequency of All Sources Coded | Percent of All Sources Coded |
|--------------------------|---|---------|----------------------------------|--------------------------------|------------------------------|
|                          | Total                                       | Percent |                                  |                                |                              |
| African American         | 28  | 35      | 9                                | 45                             | 8                            |
| Non-Hispanic White       | 17  | 14      | 6                                | 25                             | 16                           |
| American Indian (site 1) | 16  | 21      | 7                                | 20                             | 12                           |
| Alaska Native            | 11  | 9       | 8                                | 14                             | 9                            |
| Native Hawaiian          | 13  | 12      | 8                                | 20                             | 12                           |
| Asian                    | 9   | 15      | 5                                | 15                             | 9                            |
| Hispanic                 | 7   | 5       | 8                                | 16                             | 10                           |
| Generalized Site         | 2   | 4       | 3                                | 6                              | 4                            |
| <b>Total</b>             | <b>103</b>                                  |         | <b>54</b>                        | <b>161</b>                     | <b>100</b>                   |

Note: Percentages are rounded to the nearest whole number.

Source: Comparative Ethnographic Studies NRFU/UE Analysis Sample Dataset

Different patterns were observed across the sites in terms of coder/matcher coding of sources of inconsistency. In the American Indian site, the sources of inconsistency were ranked in the following order in Table 13: no official address for housing unit (n = 7), inconsistent living arrangements or incomplete interview (n = 2 each), and omission of questions, cycling for custody, and duplicate records from multiple census operations (n = 1 each). In contrast, in the Alaska Native site, interviewer error (n = 4) was the most frequent, followed by extended stay elsewhere due to work detail (n = 3), and respondent confusion, respondent ambiguity, in-mover status, and military deployment (n = 1 each).

Another pattern of inconsistency sources was seen in the Native Hawaiian site, where multiple housing structures on property (n = 5) was the most frequent inconsistency source, followed by “not included on census form, but mentioned in NRFU interview (n = 5), question rewording (n = 3), extended stay due to work detail (n = 2), and possible miscount due to cycling for college (n = 1).

The pattern in the non-Hispanic White site revealed other sources: knowledgeable proxy (n = 14) was the most frequently coded source, followed by cycling for custody and out-mover status

(n = 2 each) and question rewording and omission of question (n = 1 each). In a later section, a vignette will be presented demonstrating how one unknowledgeable proxy gave information on six persons in one household in this site that were completely different from the five who had been counted there in the census.

In sum, 247 taped interviews and 34 untaped interviews with substantial interview reconstructions or summaries were pooled into one overall dataset to analyze results across sites, while coding the remaining interviews with less complete data into a separate dataset. In the analysis dataset, 257 of the observed housing units were occupied on Census Day and 786 persons were identified. Our coder/matcher team members for NRFU and UE found 103 persons in 44 observed households with an inconsistency between the four data sources. The most frequent team sources of inconsistencies were unknowledgeable respondents, question rewording, and the omission of questions. In order to understand how these persistently undercounted groups may be differentially impacted by cultural and other factors, the discussion now turns to ethnographic reports for the groups documented.

Table 13: Specific Team Sources of Inconsistency for Persons with at Least One Inconsistency across Race/Ethnic Sites

| Specific Inconsistency Source  | Race/Ethnic Site |           |           |           |          |           |           |           | Total      |
|--|------------------|-----------|-----------|-----------|----------|-----------|-----------|-----------|------------|
|  | AA               | AI        | AN        | AS        | GS       | HI        | NH        | WH        |            |
| Unknowledgeable proxy  | 7                |           |           |           |          |           |           | 14        | 21         |
| Question rewording   | 16               |           |           |           |          |           | 3         | 1         | 20         |
| Omission of question(s)  | 10               | 1         |           |           |          | 1         | 1         | 1         | 14         |
| Person mentioned in interview but not found on Census form                   |                  |           |           | 4         |          | 1         | 3         | 5         | 13         |
| Incomplete interview   | 2                | 2         |           | 4         | 2        | 1         |           |           | 11         |
| Interviewer error: Census records does not match NRFU interview              |                  |           | 4         |           |          | 3         | 2         |           | 9          |
| Multiple housing units on property   |                  |           |           | 3         |          |           | 5         |           | 8          |
| Respondent confusion   | 2                |           | 2         |           |          | 4         |           |           | 8          |
| No official housing unit address   |                  | 7         |           |           |          |           |           |           | 7          |
| Partially inaudible tape   |                  | 6         |           |           |          |           |           |           | 6          |
| Respondent annoyance   |                  |           |           | 3         | 2        | 1         |           |           | 6          |
| Extended stay/work detail  |                  |           | 3         |           |          |           | 2         |           | 5          |
| Respondent ambiguity   |                  |           | 1         |           |          |           | 3         |           | 4          |
| Inconsistent living arrangements   | 1                | 2         | 1         |           |          |           |           |           | 4          |
| Respondent refusal to answer   | 2                |           |           |           | 2        |           |           |           | 4          |
| Interviewer makes assumptions  | 4                |           |           |           |          |           |           |           | 4          |
| Cycling for college  |                  |           |           | 1         |          | 2         | 1         |           | 4          |
| Cycling for custody  |                  | 1         |           |           |          |           |           | 2         | 3          |
| Confusion: language barrier  |                  |           |           |           |          | 3         |           |           | 3          |
| Possible miscount: outmover  |                  |           |           |           |          |           |           | 2         | 2          |
| Duplicate records due to multiple Census operations                          |                  | 1         |           |           |          |           |           |           | 1          |
| Possible miscount: inmover   |                  |           | 1         |           |          |           |           |           | 1          |
| Enumerator did not fully read questions (uncooperative, evasive Respondent)  | 1                |           |           |           |          |           |           |           | 1          |
| Military deployment  |                  |           | 1         |           |          |           |           |           | 1          |
| Not included in NRFU standard interview, but person listed in Census records |                  |           | 1         |           |          |           |           |           | 1          |
| <b>Total</b>   | <b>45</b>        | <b>20</b> | <b>14</b> | <b>15</b> | <b>6</b> | <b>16</b> | <b>20</b> | <b>25</b> | <b>161</b> |

Note: AA = African American site, AI = American Indian site, AN = Alaska Native site, AS = Asian site, GS = Generalized site, HI = Hispanic site, NH = Native Hawaiian and Other Pacific Islander site, and WH = Non-Hispanic White site.

Source: Comparative Ethnographic Studies NRFU/UE Analysis Sample Dataset

#### **5.1.4 Reasons for inconsistencies: ethnographic report cross-cutting themes**

In comparing across the ethnographic site reports, we identified themes that may affect coverage that cut across two or more of the research sites. These include: 1) enumerator error; 2) difficulty gaining access to respondents; 3) language issues; 4) cultural differences; and 5) mobility.

##### **5.1.4.1 Enumerator Error**

One aspect of enumerator error identified across all of the race/ethnic site reports was enumerator rewording, omitting, or combining of key coverage questions, despite frequent emphasis during enumerator training to read questions verbatim. Enumerator error in the form of question rewording or omission is the focus of this section; other types of enumerator error will be identified in the ethnographic profiles in the following section.

The extent of question rewording and omissions in the NRFU operation has already been documented in the earlier 2010 Census Program of Evaluations and Experiments (CPEX) evaluation, “Behavior Coding in the Nonresponse Followup Operation” (Childs and Jurgenson, 2011). That study documented that of all taped interviews analyzed enumerators read questions as worded in just 37 percent of all questions asked. Those findings are directly relevant to this report because 193 (95 percent) of the behavior-coded tapes were recorded by the ethnographers in this evaluation. This sets the context for our discussion of how and why enumerators might reword or omit questions that they were trained to read verbatim and how this may affect our key coverage questions.

To our knowledge, this is the first ethnographic study in which interviewer error has been identified as a source of possible coverage error during a decennial census. This was made possible by permitting researchers to send trained observers to accompany enumerators and systematically observe, tape, and debriefs in live census interviews over nine days across multiple sites at the same time. While observations of live field operations have been done in past censuses, most former observations were done by individuals working alone without any unifying framework of observation who observed for a few days in one site and submitted individual trip reports. While such reports are informative, they are not strictly comparable with each other and overall analysis is limited.

Question rewording and omission was a key issue in the African American site, the non-Hispanic White site, the Native Hawaiian site, the Generalized site, and the Hispanic site. Rewordings and omissions included dropping key phrases from questions, such as the critical reference date of April 1, 2010 and changing or deleting critical words such as “living” or “staying,” combining or omitting questions, failing to probe and read all response categories, and turning individual-level questions into global questions for all in the household. Ethnographers cited various reasons for enumerator questionnaire rewording and omissions:

1. Poor interviewing skills
2. To make the questionnaire conversational and put the respondent at ease

3. When enumerators were tired or rushed
4. To complete interviews quickly to be more productive per trip or day

For example, the non-Hispanic White site ethnographer documented that Question S5—the key respondent-provided household population count question documenting how many people were living or staying there on Census Day—was read correctly in only 37 percent of the interviews he observed (Albee, 2011). The most common error was to omit “April 1, 2010” from the end of the question, which unmoors the question from the Census Reference Day and might result in some people being wrongly included or excluded if the respondent thought he was being asked to report who was there at the time of the NRFU interview in mid- to late May as the reference date, instead of April 1, 2010. In the Native Hawaiian site, Question S5 was read correctly in just three of thirty interviews.

This key question, S5, as well as a case study from one site can serve as an example to demonstrate possible coverage problems that can occur when the question is misread or omitted. Major question rewording and omission of the key coverage questions S2, S5, and Roster Question 1 were the main source of inconsistencies and possible coverage error for four of seven persons rostered in one Native Hawaiian housing unit.<sup>11</sup>

In an interview at the Native Hawaiian site, the enumerator initially asked a variant of S2, “Did you or anybody live here on April first? The respondent answers, “Oh yeah, my kids and my wife.” The enumerator then skipped whole household usual home elsewhere question S3 and the lead-in statement on the residence rule. Without any reference to the List A include and exclude lists which may have helped this respondent in framing his answer, the enumerator asked his very changed version of S5:

E: “Ok, you were living here. Ok. How many people live in the house all together? Your wife...”

In this very altered version of question S5, he has both omitted the reference to April 1, 2010 and changed the verb tense from “were living” to the present tense, implying he is now asking about who is living there on the Interview Day in May, rather than on April 1, 2010. Note also that “or staying” has been dropped, and that “all together” has been added by the enumerator. The enumerator then truncated the roster question and simply asked for names, again without reference to April 1, 2010. The respondent identified his wife and five children for a total of seven in the household counted on the EQ form. However, based on the ethnographer’s later phone followup, four of the seven people should not have been counted there and represent possible overcounts.

While most of the question rewording or omissions were situational and represented enumerator efforts to complete difficult interviews as well as possible, a very small number of enumerators appeared to be deliberately shortening or combining questions in order to complete more

---

<sup>11</sup> The actual wording, formatting, and placement of this question on the EQ can be seen in Appendix C

interviews in a given area or time period and be recognized for their high productivity. In the Hispanic site, the ethnographer said that enumerators felt they were under pressure to produce as many interviews as possible and this may have led some to cut corners in question wording. This was also documented in the Generalized site and the Alaska Native site. The implication, at least for this small number of enumerators, is that the enumerator incentive structure may be a bit more oriented toward quantity as compared to quality.

These forms of interviewer error can also be placed in a larger context. In questionnaire administration, there is tension between enacting procedures to get valid and reliable data and having the flexibility to: (a) operate effectively in the field, (b) engage a wide variety of respondents, and (c) get enough information to complete the interview. This is an ongoing discussion in survey methodology.

Frequent major changes to questions documented in the behavior coding study and in this section are a concern because they may affect the accuracy, completeness, validity, reliability and comparability of census data. Suggestions are offered later for addressing this issue in training, in questionnaire design and testing, and in more research on interviewer/respondent interaction.

#### 5.1.4.2 Difficulty in Gaining Access to Respondents

Another theme common among site reports is the difficulty enumerators and ethnographers had with accessing respondents in the form of locked buildings and housing units located in rural areas. These issues were reported in the Asian site, the African American site, the American Indian site, the Native Hawaiian site, and the Alaska Native site.

Difficulties in accessing respondents took a diversity of forms. In the African American site and the Asian site, this took the form of difficulty in accessing respondents in locked or otherwise inaccessible buildings. At the African American site, many of the apartment buildings that enumerators visited were gated or had locked doors with buzzers, making it hard to access respondents. In the Asian site based in San Francisco Chinatown, the enumerators consistently had problems with building accessibility, particularly in single-room-only (SRO) buildings, also known as *san fong*, with single rooms and shared bathrooms and kitchens, with buzzer boxes that were not working and with mislabeled mailboxes. A number of respondents at the Asian site claimed to have sent in their mailout census forms and were confused, upset, or angry at being contacted by enumerators to do the forms again. The situation with mail left in piles for residents to sort through provides a plausible reason that could explain why at least some of the persons in these buildings were being contacted in NRFU even though they claimed to have sent in their completed census forms. These situations have been called “apartment building effect errors” and they can lead to coverage complications, errors, and followups.

Difficulty in gaining access takes a different form in rural areas. As seen in the earlier section, “no official address for housing unit” was identified as a source of error seven times in both the Native Hawaiian site and American Indian site. In the Native Hawaiian and Indian reservation sites, the lack of accessibility was due to being in more remote areas lacking street numbers and addresses and having difficulty in trying to locate and verify the correct housing units.

In the Native Hawaiian site, the ethnographer noted that 14 of the housing units she saw had vague addresses with no house number posted and one was just a lot number. She mentioned there was a chance that they may have visited some dwellings that did not match the addresses in the enumerator's book, in other words, interviewing the wrong persons. On the American Indian reservation, the ethnographer reported difficulty in trying to canvass a remote area far off paved roads that took up the greater part of an afternoon. He also reported difficulty in using maps with few or no features to try to find and verify housing units in the remote areas and also in distinguishing them from other housing units in more settled areas.

A related factor on the American Indian reservation was that few of the enumerators hired on this reservation had four-wheel drive vehicles. Enumerators were held off from being sent to those remote locations until the last weeks of enumeration in May when it was hoped that the local census office would receive approval to rent four-wheel drive vehicles to enumerate those areas. If the request was not approved, or not granted in enough time, some of the most remote housing units on this reservation may have been missed in the 2010 Census. The final 2010 national CCM estimates showed an undercount of 4.88 percent on Indian reservations (U.S. Census Bureau, May 22, 2012 press release), the highest among the race/ethnic groups and sites cited. On this reservation, the geographical, logistic and cost difficulties of gaining access to persons living in remote areas may have been factors affecting coverage, if enumerators were unable to reach all of the remote housing units by the end of enumeration.

#### 5.1.4.3 Language Issues

In this section, we document issues that enumerators faced when interviewing persons who speak a variety of languages other than English and how enumerators dealt with them. Within the large theme of language barriers, we highlight problems that arose from the lack of translated enumerator questionnaires and Information Sheets, the use of on-the-fly translation, and dialect variation within a larger language. These issues occurred the most at the Asian site, the Hispanics site, the American Indian reservation, the Generalized site and the Alaska Native site.

In terms of interview language, 222 (79 percent) were in English, 23 (8 percent) were fully or partially in a Chinese language or dialect, 16 (6 percent) were fully or partially in Spanish, seven (3 percent) were partially in English and an American Indian language, and one partially in English and Tagalog (< 1 percent). Of the 16 in Spanish, nine were conducted entirely in Spanish and the remaining seven in a combination of English and Spanish.<sup>12</sup>

A lack of translated materials was an issue at the Asian site and the Hispanic site. At these sites, there were no available Chinese- and Spanish-language versions, respectively, of the Enumerator Questionnaire. However, Spanish- and Chinese-language versions of the Information Sheet were available, and provided some assistance. With no standardized Spanish- and Chinese-language version of the EQ form, enumerators had to do on-the-fly translations for all interviews conducted in these languages.

---

<sup>12</sup> These figures refer to the analysis dataset, which contains fewer cases than the total number of observed ethnographer interviews.

At the Hispanic site, nearly all of the Spanish interview translations were done on-the-fly. Bilingual and monolingual respondents interpreted the Spanish questions as they understood them. The bilingual interviewers at the Hispanic site were observed to be able to communicate for the most part but were not professional translators with the necessary skills and backgrounds to do quality translations. They had problems reading the Spanish verbatim from the Information Sheet. They also did not fully understand the respondents' answers, including being able to distinguish middle names from family names, and trying to identify family and other relations in the household. At the Asian site, on-the-fly translations were found to be tiresome, time-consuming, and difficult to do consistently. These translation issues likely resulted in major question changes, inconsistent question translation by interviewers, and inconsistent question interpretation by respondents.

Language barriers also resulted in other persons in the housing unit besides the respondent serving as an on-the-fly translator between the interviewer and respondent. For example, at the Hispanic site, of the 40 total observed interviews at the Hispanic site, 23 were conducted with Hispanics: 21 of them with the initial household respondent alone and two with a second person in the household serving as interpreter. In these two interviews, children under age 12 served as the translators, and while they did a good job of translating, there was a question of whether the scenario was appropriate. The ethnographer noted that relationships can be sensitive, especially in blended families where children might have different fathers but not know it; in such cases a father might not give the actual relationship to avoid revealing to a child that he may not be a biological son or sibling.

Another type of language issue was the variation of dialects spoken within a major language, which was only a factor at the Asian site in San Francisco Chinatown. Of the 13 interviewers observed at the Asian site, all but one spoke a Chinese language or dialect and all but one lived in or near Chinatown. The distribution of Chinese interviews by Chinese dialect is more complex with interviews fully or partially in Cantonese, Toisanese, Mandarin, and Fujian. Issues arose when the enumerator and the respondent did not speak the same Chinese dialect, which occurred in 33 of the 56 Asian site interviews that were in some type of Chinese dialect, but with no English (Lee Tu, 2011). Most Cantonese and Toisanese speakers can understand each other if they speak slowly and take time to think, but Mandarin and Cantonese speakers cannot understand each other's languages unless they are bilingual (Chen, 1999). It is not known how much the use of mixed dialects affected data quality.

#### 5.1.4.4 Cultural Issues

Cultural issues, including cultural competency, complex household networks, and a mistrust of government were observed in this evaluation. These issues were identified in the Native Hawaiian site, the American Indian site, the Alaska Native site, the Hispanic site, and the African American site.

The need for cultural competency was an issue at the Native Hawaiian site and the American Indian site. At the Native Hawaiian site, it was important to hire Native Hawaiians and Other Pacific Islanders to conduct interviews in their communities. Also helpful was the hiring of a well-known Hawaiian to be a cultural facilitator, who went into communities and introduced the



enumerators to help gain acceptance by local Native Hawaiians. In some areas, local people are not open to outsiders coming into their areas to do interviews. On Hawaii Home Lands, it may be considered trespassing for outsiders to come in and knock on doors. In the American Indian site, all of the enumerators observed were of that tribe and lived on the reservation. That was absolutely essential for gaining the trust of respondents and sometimes, for speaking in the local language with older respondents.

Complex household networks and the housing units they occupy were issues at the Hispanic site and the African American site. The prevalence of laterally extended households, with cousins, nieces/nephews, aunts/uncles, and siblings-in-law, was an issue at the Hispanic site. This may be a cultural practice, particularly if the immigrants are from specific Mexican and Central American regions that share a tradition of married siblings living in family compounds. In those regions, the families live in their own structures on the same parcels of land and may share kitchens, dining rooms, living areas, and other common areas (Garcia, 2011). Laterally extended household structures may also be a result of shortages of multi-unit housing units in some areas of Dallas that may result in family members sharing a single-family housing unit (Garcia, 2011).

The importance of “family buildings” and understanding the dynamics of extended family networks that go beyond single households was an issue at the African American site. A “family building” is a home or building that is owned by an African American, usually an older woman that is open for family members to move in and out of as the need arises. One family building in that site had an elderly mother and her adult daughter live in the upstairs housing unit and another daughter and her children live downstairs. While these are separate housing units for 2010 Census purposes, there is sharing of resources and childcare between these housing units. She also identifies another situation of a family that “lives apart together.” For more than 30 years, the woman and her children lived in one housing unit of an apartment building and her partner (the children’s father) has lived in another. They consider themselves a family but live apart because there would be too much conflict if they lived under the same roof. There were also situations with “linked households” in which children are shared between parents and among grandparents.

With current methods of enumerating persons in physical housing units in the 2010 Census and in most surveys, the Census Bureau does not have a way to identify and learn about these “social households without walls” that share resources and strong family ties and may help us to better document and understand erroneous enumerations across households. Demographers identified the collection of cross-household ties as a major issue for future family research at the 2011 Counting Couples, Counting Families Conference (Brown and Manning, 2011) and some are developing new questions for surveys to address this (Brandon, 2008; Brandon, personal communication: August, 2012).

Fear and mistrust of the government were particular issues in the Hispanic site. Due to heightened concern at that time about the new Arizona law that was seen to be cracking down on illegal immigrants, there was growing mistrust and fear among Hispanics, especially among those from other countries who had doubts that giving information to the Census Bureau would be secure from government immigration authorities. In some cases, at the beginning, the ethnographer spoke in Spanish to the respondents to reassure them the interview was safe and to

explain the purposes of the taping. He was concerned that if he did not say something during the interview, some respondent might tell others in the neighborhood about being tape recorded and then community sentiment might build to avoid participation in the 2010 Census.

#### 5.1.4.5 Mobility

Mobility was a factor in a number of records check inconsistencies across the race/ethnic observation sites. Permanent moves from one place to another or movement among primary and vacation homes and inconsistencies in recording them were identified in the Generalized site. Cycling for long-distance work was a major factor in inconsistencies and possible coverage errors at the Alaska Native site and other sites, and was often combined with other factors. Cyclical mobility in the form of children moving between households of parents, grandparents, and others, girlfriends/boyfriends moving between their respective residences, college students going back and forth between college and home residences, and snowbirds moving north and south seasonally pose challenges for enumeration in the right place. The theme of cycling for work or other reasons between residences and distant locales, together with uncertainty as to where persons should be counted, has been discussed in regard to the Living Situation Survey (Bates and Gerber, 1998), and in small-scale qualitative studies on the Navajo reservation, among the Inupiat, Hispanics, Koreans, African Americans and non-Hispanic Whites (Tongue, 2006; Schwede, 2009; Craver, 2006; Goerman, 2006; 2006; Kang, 2006; Holmes 2006; Childs, 2006; Blumberg 2006). Summaries of some cases that were inconsistent due to mobility issues are provided in the next section.

### 5.1.5 Ethnographic site profiles: factors contributing to identified record inconsistencies

In the ethnographic site reports, there were also characteristics pertaining to specific sites that are factors in inconsistencies and possible coverage errors. This section provides profiles of each research site, a discussion of factors affecting inconsistencies as well as summaries of actual interviews to illustrate the types and sources of inconsistencies and possible coverage errors that were found. Some of these may be of use in reviewing current questionnaire materials and procedures and suggesting improvements for the next census in 2020.

#### 5.1.5.1 The Alaska Native Site (Bunten, 2011)

This borough includes Kodiak Island, nearby islands, and part of a peninsula. Accessible only by plane or boat, the large island only has around 100 miles of paved roads and vast rural areas. The small city has large warehouses, a marina, a fish processing plant and a boatyard, as well as a Coast Guard base and a semi-permanent work force. The borough has a complex maritime economy. Alaska Natives are just over one-eighth of the Borough population and live dispersed among the population, making them hard to find. In addition to Alaska Natives, Whites, Filipinos and others were also in this observation study.

In this site, mobility for work and cycling was a major factor in inconsistencies and possible coverage errors, often combined with other factors. A number of the respondents or their family members were deployed in the Coast Guard or in the Merchant Marines and often out at sea. In a few cases of such mobility, interviewer omission of the April 1, 2010 reference date led to

husbands being included on household rosters who likely should have been counted elsewhere: one man who had been deployed to a distant city and another man who had been out on a Merchant Marine vessel for more than half the year leading up to and including April 1, 2010. These cases are further complicated because these men may have been included on individual census forms in the group quarters operations. While most of such housing unit/group quarters duplications would be identified and resolved in unduplication operations, there is a chance some would remain in the count as unidentified duplications.

The intersection of mobility between a household and a worker's dormitory (another type of group quarters), with interviewer error and one man's attempts to avoid being counted were highlighted in another case. A man present during an interview with his father was not included in the household count initially because he was not there on April 1, 2010. When asked if he had been counted in the census, the son replied, "That's how I skated out of answering any questions. I told them I wasn't here!" The son had been away working on the North Slope oil fields for around five months and there on April 1, 2010. It is not clear if he should have been counted at that worker's dormitory on April 1, 2010 or here because he comes here when not there. If, at the dormitory, he avoided filling out an individual census form that includes a line for usual residence elsewhere, he could have been counted at the dormitory and also at his father's house, causing a potential duplication. Another such case will be discussed in Section 5.2.

Mobility in the form of short- and long-distance cycling for non-work reasons was also seen. In one case, this combined with possible respondent concealment and interviewer intrusiveness when a respondent did not list his girlfriend. Knowing the respondent's living situation, the enumerator brought up her name and probed on whether she should be counted here or at her parents' house. Somewhat reticent, the respondent said she was staying with her parents and was likely counted there. It is not clear he would know that.

In the final example, seasonal mobility of retired snowbirds combined with interviewer error in counting the man here when he clearly stated he had submitted a mailout form for his usual residence in a southern state. While it did not happen in this case, it may be helpful to future coverage studies in Alaska to consider that some snowbirds might claim their Alaska address out of concern that their annual government dividend checks might be jeopardized if they were not counted in Alaska.

In this site, factors in inconsistencies and possible miscounting of some included multiple types of interviewer error, mobility for work and housing unit/group quarter crossovers, local and inter-state cycling among households, and a man's attempts to avoid the census.

#### 5.1.5.2 The American Indian Site (Lahren, 2011)

This is a remote, moderate-sized reservation in the Southwest with more than 90 percent of its residents reported as American Indian. Agriculture is the basis of this group's life and the production and sale of arts, crafts, and jewelry provides sources of income for some. Due to a shortage of formal sector jobs on the reservation, many go off-reservation to work. There is a strong cultural identity on this reservation, with many speaking a native language. More than half on this reservation spoke a language other than English at home and less than one-quarter

speaking English “less than very well” (2006-2010 American Community Survey American Indian and Alaska Native Tables). Religious practices and ceremonies are also an integral part of this identity.

There were a number of settlement areas as well as individual houses scattered widely through remote rural areas. Many houses were on unpaved roads with no street names or addresses and could only be reached by four-wheel drive vehicles. Assignment area maps had few, if any distinguishing landform features, lacked topographic details and were not uniform in scale. They often had just physical descriptions of housing units, such as white trailer with gray trim.” All of these factors made it difficult to find some mapspotted houses and to verify they were correct.

The lack of a street name or number and census identification number appeared to be primary factors in inconsistencies and possible coverage errors in several cases. In one of these, an elderly woman said a grandson had been staying with her but moved out to stay with his girlfriend, without providing his name, alternative address or move date. They both appeared to be duplicated: she, in one house with a descriptive address and in another with no address, and he, in one place with her and also at a different place. Without a firm address and move date, it was not clear where to count the grandson or which record was correct for the respondent. In another case, the interview strongly suggested the persons should be counted there but they were not found in the final localized 2010 Census unedited dataset clusters available for this site. It is not clear if this is due to the limited geographical dataset or if they are in a different cluster and did not show up, or were missed.

In a different interview, enumerator familiarity with the respondent and her family may have led her to skip key coverage questions on whether they had all lived there on April 1, 2010 and simply list the persons she knew to be there. Later the respondent stated that one daughter was staying elsewhere at her father’s house through the school year to attend a different school with more services. The daughter was in the final localized 2010 Census unedited dataset for this housing unit when she should have been counted at her father’s house. There is a possible duplication if she was also counted at his place. It is not known in another interview why only the fifth of six persons listed then could not be found in the final localized 2010 Census unedited dataset.

On the positive side, potential language and cultural barriers were mitigated because all enumerators were tribal members who were familiar with the language and culture and who had pre-existing relationships with some respondents. Some interviews were conducted partially and one wholly in an Indian language via on-the-fly translations, since there was no language-specific form for this group. While this resulted in getting answers to all questions, it likely resulted in variation in question wording translations.

For this American Indian site, the lack of street names and numbers appeared to lead to possible unresolved duplications in one case and possible omissions in another. Interviewer error in omitting key questions was a factor in counting a daughter in the wrong place. Having local tribal members conduct interviews overcame potential language and cultural barriers, but also may have resulted in variation in question wording as interviews were translated on the fly.

### 5.1.5.3 The Native Hawaiian Site (McMillen-Wolfe, 2011)

The Native Hawaiian site located in Hawaii County, also known as the Big Island, included Hawaiian Home Lands and general population areas. Hawaiian Home Lands are specially designated areas reserved for ownership by Native Hawaiians who are at least 50 percent blood of races that lived in the Hawaiian Islands before 1778. Gaining access to Hawaiian Home Lands can be difficult and in some areas residents do not accept outsiders. Housing shortages and landlessness are large problems. Native Hawaiians are disproportionately affected by poverty and disproportionately represented as victims and offenders, making them harder to enumerate. They are also much more likely to live in large households (McMillen-Wolfe 2011), also putting them among hard-to-count populations.

As in the American Indian site, this ethnographer reported vague addresses and inadequate housing descriptions. Some rural census maps had insufficient information, making it difficult for enumerators to orient themselves geographically. When map spots were not in numerical order, it was difficult to locate addresses. Vague and contradictory addresses were a problem in five housing units in a relatively new condominium community, where the enumerators only had mailing addresses, such as 123 Hawai'i Place, but the street address was different: 45-678 Big Island Drive, Unit 333 (not the real addresses).

The ethnographer noted four cases of possible hidden housing units. In one, seven persons were rostered. The ethnographer's question about a baby in the room who was not included on the roster led to the identification of two other "hidden" housing units on this large rural property with dense foliage. In the first hidden housing unit, a man thought to be living there was not present; a notice of visit was left. At the second "hidden" house, five people, including the baby, were listed in the observed interview. However, in the consistency check across data sources, of the total of 13 persons on this one property, just five of the seven in the main house were listed in the final localized 2010 Census unedited dataset. One possible explanation is that the continuation form for Person Six and Person 7 in the main house may not have been processed. The family and the man in the two discovered housing units were not in the final localized 2010 Census dataset; it appeared that just the one housing unit on this property was listed. In another case, a separate housing unit in the back of another house was rented, but his name was not obtained and he was not in, so that was not an observed case. In another, the ethnographer noticed a possible separate attached unit, but did not ask and the enumerator seemed not to notice it. It is not clear if that was a separate, possibly occupied unit.

A case of respondent attempts at concealment and avoidance occurred when the respondent was identifying just herself on the roster. Two men standing nearby overheard this and one urged her to include the other man as living here. Not pleased, the woman did provide his name in question H1 and then said she herself was not there on Census Day but at another place, which she declined to identify. In the final localized 2010 Census unedited dataset, the man identified in question H1 was the sole resident; the original respondent was not found, perhaps because she had not given the alternative address. The man seemed to be an employee who also had some ties to her other place; she seemed to not want to be counted with him.

In summary, the Native Hawaiian site had vague addresses in both a rural and a more urban area as well as residents in two hidden housing units that may have been missed in the 2010 Census. In another case, the H1 omission flag question identified a man who was counted there in the final localized 2010 Census unedited dataset, even though the respondent herself was not counted there after she had listed herself as the sole resident.

#### 5.1.5.4 The Hispanic site (Garcia, 2011)

The city of Dallas has a large and rapidly growing Hispanic population; as of 2010, 42 percent were Hispanic (Garcia, 2012). Parts of this site were mailed the special English/Spanish “swim lane” form to make it easier for Spanish-speakers to complete.

This ethnographer stated that passage of anti-immigrant legislation that ignited a national debate on immigration just before the NRFU operation was a likely factor in the large 10 percent soft refusal rate he observed. That Arizona law, combined with ordinances in two Dallas area cities aimed at identifying illegal immigrants through police stops or reporting of immigration status when renting apartments, likely heightened Hispanic concerns if they answered census questions. Other Hispanics complied but were not fully forthcoming, such as claiming not to know birthdates; the ethnographer thought they might be concealing data in fear of being fully identified and deported (Garcia, 2011). Immigration status and passive and active resistance with strangers are factors he identified in Hispanic undercounts in prior research, along with limited or no knowledge of U.S. censuses, language and illiteracy, irregular housing and household arrangements, and mobility (Garcia, 1992 and 1995; Garcia et al., 2008). Other factors he identified in this site include: enumerator question wording shortcuts to complete as many interviews per trip as possible; limited resources; the lack of a Spanish-language questionnaire; non-Spanish speaking enumerators; and illiterate respondents not able to use the Information Sheet.

In one interview, the language barrier, household complexity of three families living in one housing unit, and a respondent who was not the owner and may not have been knowledgeable or forthcoming and unknown factors resulted in three inconsistencies with the final localized 2010 Census unedited dataset. Because he was unemployed, the man and his family were staying with his sister. When asked how many people lived here, he answered, “Many” and was disappointed to be asked for all names, giving nine, with no cue of anyone else. However, the coder found three more people here in the final localized 2010 Census unedited dataset. It is not clear whether the three were indeed living here and the respondent was just unknowledgeable or whether he concealed information or something else. This is one of the inconsistent cases in which new persons appear in the final localized 2010 Census unedited dataset records and the authors do not know why. The non-Spanish-speaking enumerator had trouble recording data and many discrepancies were found in the names and birthdates needed for matching and unduplication.

In an example of possible concealment, a woman sitting in a luxury car in front of a rundown house completed the interview, reporting her family as living there. Several cars parked on the property suggested the type of place where single rooms in dilapidated houses are (illegally) rented to transnational migrants. The garage appeared to be used as sleeping quarters. The ethnographer suspected this woman and her family lived elsewhere and rented bedrooms (and

maybe the garage) to migrants, who may or may not be documented, which is common. This potentially provides insight into an underground economy of substandard housing for immigrants in a place where rentals are few and expensive.

Finally, in two households, infants present during the interview were not included in the household count or roster, but were identified in question H1 or in a debriefing. In one, a Black woman mentioned the child in omission question H1 and later it was learned he was a cousin who lived there on April 1, 2010. In the other, the Hispanic woman did not include her one-year old child on her lap. It is not known if these names were entered in question H1. These children are possible omissions who were not in the localized final dataset. Past studies indicate that children aged 0 to 4 are persistently undercounted (Mule, 2012; O'Hare, 2013).

#### 5.1.5.5 Asian, primarily Chinese site (Lee Tu, 2011)

The Asian, primarily Chinese, site was in San Francisco Chinatown, the oldest and second largest concentration of Chinese in one U.S. community. Housing, where many low-income Chinese families live, is in a large project, or in Single-Room Occupancy (SRO) hotel units known as *san fong*, or in regional benevolent association buildings rental rooms. Chinatown is comprised mostly of low-income elderly and recent refugee/immigrant families. Many speak Cantonese and Toisanese and more are speaking Mandarin; in Census 2000, 53 percent spoke English not well or at all (Zinzius, 2005). Cantonese and Mandarin are mutually incomprehensible; they share the same writing system but speak the words differently. Cantonese and Toisanese speakers can understand each other if they speak slowly and concentrate.

Gaining access to Chinatown respondents was a major issue; one-third of the interviews were in buildings with vague or no exterior signs. All of the SRO *san fong* buildings had disabled buzzer boxes and mislabeled mailboxes, sometimes done deliberately by tenants for security reasons. Mail was often left in piles for residents to sort through.

Language was a big issue. All but one enumerator spoke a Chinese language (mostly Cantonese) and/or dialect and lived in or near Chinatown. Of the 62 observed interviews, 59 were in one or more Chinese languages or dialects. Of those 59 interviews, enumerators translated Cantonese on the fly and respondents answered in Toisanese in 23 interviews.

The ethnographer mentioned bilocality—splitting residential time between two places—linked to transnational movement, kinship ties, and/or lack of sustainable Chinatown jobs—as factors in possible coverage error. Bilocality, kinship, and transnational mobility were factors in one inconsistency where the respondent mentioned her mother in possible omission question H1. The mother had lived there for more than a year but returned to China after April 1, 2010 when her visa expired. She should have been counted here as an outmover but was not found in the final localized 2010 Census dataset.

That same interview had more inconsistencies. Language was a factor, for she did not know the official spelling transliteration of her husband's and child's Chinese names in English. The respondent gave the enumerator a document with the names and demographic characteristics in

English which the enumerator silently copied onto the EQ form. Lacking the names and demographic characteristics on the tape and transcript, the coders could not definitively match the husband and child to persons in the final localized 2010 Census unedited dataset. (Five other NRFU respondents also found documents to provide the official English spellings).

Two more persons were also identified in question H1 in this same interview, but the names of the unrelated woman with a child who was renting the back room of the apartment and sharing the kitchen were not provided. It seemed that the renter did not have a separate outside exit, making this an irregular housing situation in which two unrelated families live separately in one housing unit but are classified as one household. By not including these two in her household count or roster, the respondent showed she did not consider them part of her household. The factor is a mismatch of the respondent's *emic* view of who belongs in her household with the 2010 Census definition of those sharing a physical housing unit, regardless of whether they function as a household. This may lead to omissions in doubled-up households with boarders and backroom renters. Other persons were found in the final localized 2010 Census dataset for this housing unit, but without identifiers, the persons alluded to in the interview cannot be definitively matched.

Another interview also raised the question of what a housing unit is, but in the opposite direction. Three persons used three contiguous *san fong* units as one home, but each unit was on the 2010 Census NRFU address list as a separate SRO housing unit. The enumerator tried to complete forms for unit 1 and unit 3, but the respondent insisted there was just one housing unit comprised of the three rooms: one for sleeping, one for a living room and one for other purposes. How should this be resolved? Unit 1 and unit 3 are being used as if part of one housing unit, but not to sleep in, but they are set up as separate units for rental, and were listed as three separate housing units on the Master Address File. Was this an ad hoc use of unoccupied units or did they rent all three rooms formally? The respondent says it is one housing unit, so should just one form be completed and unit 1 and unit 3 be deleted from the MAF? Or should unit 2 be occupied by the end units be classified as vacant or nonresidential? The enumerator appeared to record them as living in unit 2, with unit 1 and unit 3 vacant, but that may not have been right. The respondent was angry, claiming she had already mailed in two census forms, so why was the enumerator bothering them? She could have completed mailout forms in Chinese, which may not have been processed by the April 19, 2010 NRFU cutoff date. Or she may have filled out forms barcoded for other units, with no return processed for her unit. The irony is that after all this, these three persons were not found in the final localized 2010 Census unedited dataset and may have been missed.

Other cultural factors that may indirectly affect coverage by complicating matching and unduplication efforts: 1) some providing birthdates in Chinese lunar years, just 354 days long; 2) Chinese naming customs variation (half of the respondents structured their names as last/first/middle name while the rest gave them as first/middle/last name); and 3) many wives not taking their husband's family names, though their children do.

In this Asian site, it appears that vague addresses, language and spelling barriers, bilocality, listing problems, doubling up, use of multiple rooms as a single housing unit, mismatches in respondents' and official definitions of "household" and "housing unit," apartment building



effects, and the possibility of slow processing of forms in languages other than English may have been factors in inconsistencies.

#### 5.1.5.6 African American Site (Jarrett, 2010)

The African American site consisted of three Chicago neighborhoods. One neighborhood was ethnically diverse and had both a strong middle-class presence and also areas where poorer residents lived with more crime. Apartment buildings of three flats to multiple stories predominated but some blocks had large homes. The second and third communities were more than 90 percent Black and impoverished, with around 40 percent to 50 percent below the poverty line in Census 2000. Around 80 percent to 90 percent of the housing was in mostly multi-story apartment buildings. One of these had some single-family houses and the other had boarded-up buildings and vacant lots. Poverty and crime were issues.

Access to respondents was a key factor. Many buildings had iron gates and enumerators had to be buzzed into the buildings. Often doorbells were unlabeled, making it difficult to connect names and units. Some building managers were helpful while others avoided enumerators.

Some observed housing units and persons in them seem to have been missed in the 2010 Census. The records check showed that three of the observed addresses appeared at first to have completely different households on Census Day, referred to as a “Smith/Jones” outcome. In one, the ethnographer said the respondent usually lived in the English basement housing unit, was there on Census Day and should be counted there. Just one woman, however, with a different name and years younger was found at this address in the final localized 2010 Census unedited dataset records. The team reviewed this and decided the English basement may have been missed while the main part of the house was included. The same issue may explain inconsistencies in another where the ethnographer said three persons should be counted in the basement of a three-floor house, but the final localized 2010 Census unedited dataset recorded two different persons at the address. In one case, the respondent was a man in a senior home, but a different person was in the final localized 2010 Census dataset. It is not clear if this was a missed unit miss or if the respondent had moved there from another room since April 1, 2010; respondent confusion from cognitive impairment was also a factor. In these cases, coders searched these persons in the final localized 2010 Census unedited dataset but did not find them.

Enumerator error in not referencing April 1, 2010 for all persons led to inconsistencies when the coder team could not determine if the persons should be counted there but the final localized 2010 Census unedited dataset did count them there. In one case, the enumerator asked the teenage respondent if she had lived there on April 1, 2010 and she said yes. The interviewer then asked a very truncated version of question S5, “How many people are in your household?”, and listed five more. Also, the interviewer inverted overcount question 7 by asking “Does anyone sometimes live *here* for any of these reasons...” which precluded learning if anyone had *another* place. The team coded these as “unable to determine where to count,” which was inconsistent with the final outcome. Although it may have been the right outcome, the enumerator’s major question rewording did not provide enough evidence to support this.

Respondent hostility and concealment were factors in the inconsistency in one case. The respondent said she moved here in March and sometimes stayed at her sister's but refused to give that address or say how much time she spends at each place. She clearly stated she did not want the government to know where she lives. Without these data, the ethnographer and team decided this was unresolved and she could be a possible unresolved duplication across housing units, but the final localized 2010 Census unedited file did count her in this location.

Mobility and tenuousness, distrust, possible concealment, confidentiality concerns, and fear of adverse consequences were factors in another case. A young adult son on probation answered the enumerator questions while consulting with his girlfriend by cell phone about this situation. He did not include himself, saying he did not live there, but stayed with relatives and friends without giving any details. He appeared to have stayed overnight there and just gotten up; the ethnographer could not decide if he should be counted there. The team coded him as out-of-scope because he said he did not live there, but it is possible he deliberately left himself out.

In the African American site, multiple factors figured in inconsistencies: enumerator error, mobility, possibly missed housing units, respondent concealment and hostility, mistrust of government, and confidentiality concerns. In some cases of enumerator error, the persons were counted in the 2010 Census, despite adequate within-interview substantiation.

#### 5.1.5.7 Non-Hispanic White Site (Albee, 2011)

Independence, Missouri, a satellite of Kansas City, is the fourth largest Missouri city, located primarily in Jackson County but also in Clay County. This county was estimated to be 63 percent White, 24 percent Black and the rest in other races (2011 Census Populations Estimates program). In summer 2010, the effects of the national recession lingered; the number of persons in families in poverty had more than doubled in one year (ACS Trend Report: Jackson County, Missouri, 2008-9). Independence had a mix of well-maintained homes in some areas and homes in serious disrepair, sometimes abandonment, in others.

One "Smith/Jones" case accounts for a disproportionate 11 inconsistent persons. This interview was conducted with a young African American proxy who said he was visiting in this lower-income Section 8 housing project almost exclusively inhabited by African Americans. Unlike other cases, this one was done during a blitz enumeration in which many enumerators, crew leaders, and managers stayed close to each other and enumerated in pairs. None appeared to be African American. The respondent said he was visiting a mutual acquaintance and did not know much about the people there. He named two men and mentioned four biracial stepchildren, but did not know their demographic characteristics. None of these six was found in the final localized 2010 Census unedited dataset in this housing unit or elsewhere. Rather, a Black mother and her four children were listed. While it is possible he was just unknowledgeable, it is plausible he just made up information about the men and children to comply with the 2010 Census while concealing real information during this blitz enumeration with so many persons there to collect data for the government. Section 8 housing often has strict limits on how many persons can stay in a housing unit without risking eviction. The coder decided that the data from this interview were likely wrong and that the final 2010 Census dataset counted the right persons here. This housing unit was likely reinterviewed later as a result of a crew leader invalidating this

interview because proxy respondents are not allowed in first interviews. Note that all 11 persons in this household were classified as having inconsistencies, including those five who were assessed to be the Census Day residents here. This one case accounts for many of the unknowledgeable proxies in Table 13. It also demonstrates that persons with inconsistencies across data sources are not necessarily persons with coverage errors.

Another “Smith/Jones” type of inconsistency involved a respondent who clearly stated that he lived in this house alone since his son moved out last summer. The ethnographer verified this in a call to get answers to correctly worded questions. In the local final localized 2010 Census unedited dataset, however, two adult siblings were listed by a proxy respondent—this man was not found. Due to the proxy, the coder classified the respondent as correct and the siblings as “unable to determine where to count.” Data from proxy respondents have been shown to be less complete than those from household members (Compton and Bentley, 2012). There was no mention of any hidden housing unit.

Mobility in the form of children cycling back and forth among housing units was a factor in two interviews. In one, the respondent mentioned in H1 that his stepson lives elsewhere but stays here every other weekend and every other Wednesday. A later CFU interview was conducted; this stepson was included in the final localized 2010 Census unedited dataset. As mentioned, the NRFU H1 question does not establish where someone identified in this question should be counted, but the CFU does, and this case shows the outcome change.

In the second case, a godchild picked up in the H1 omission question was said to be here on April 1, 2010 and from Thursday to Sunday nights. The ethnographer counted her, but the coder said she spent four nights and more full days away. Mobility that is almost evenly divided between two places is one of the most difficult living situations to resolve. The child was correctly not counted here in the final localized 2010 Census unedited dataset.

In the non-Hispanic White site, mobility in the form of children cycling between housing units accounted for two inconsistencies across data sources. In one, the cyclist was counted there, in the other she was not. Cycling cases can be very difficult to resolve. In the other case, a proxy was either unable to present accurate information or may have deliberately provided incomplete and perhaps false information but the real situation cannot be determined.

#### 5.1.5.8 The Generalized Site (Blumberg, 2011)

Broward County, Florida changed markedly from a majority non-Hispanic White area in Census 2000 to a majority minority area in 2010. Non-Hispanic Whites of one race fell from about 58 percent to 44 percent, while African Americans increased from 21 percent to 26 percent, and Hispanics with any race alone or in combination grew from about 17 percent to about 25 percent, with the remainder spread across other groups ((Blumberg 2011: 16: American FactFinder, 2010 Census, SF-1).The county economy relies heavily on tourism and on snowbirds that had already begun leaving by Census Day and were gone by the mid-May NRFU interviews. This area was very hard hit by the economic crisis and unemployment and many housing units were vacant due to foreclosure or abandonment or to snowbirds having left.

High vacancy and turnover were reflected in the interview attempts here. A number of the ethnographer's observed housing units were reported to have no population on Census Day because they were vacant, were seasonal or second housing units with a whole household usual home elsewhere (WHUHE), or had been converted to nonresidential use.

Records check inconsistencies in this site were almost all due to people who had moved in after April 1, 2010 ("inmovers") who should have been counted at their prior April 1, 2010 places. They were not part of the analysis sample due to our focus on observed housing units. But what happened in this site is instructive in showing possible error at prior April 1, 2010 places. For in-mover and WHUHE housing units, the enumerator manual (but not the questionnaire) stipulates that enumerators take added steps after completing the observed address EQ. They were to ask in-movers if they had completed a census form for their prior address. If yes, the persons had been counted; nothing more was needed. If the in-mover said no, or did not know, the enumerator was to complete a new form for the prior address.

The ethnographer documented that this procedure was followed in just one of three in-mover interviews and one of two WHUHE interviews (Blumberg, 2011). It is thus not known whether in-movers in the remaining three interviews were counted at their previous places in the 2010 Census. In one of these, the failure to ask was critical, because the respondent and her son moved into the new housing unit right on April 1, 2010 itself, when the risk of duplicating persons across former and new housing units is highest. In the second of those, the interview was conducted by an enumerator admired for completing more interviews per day than almost all others; unfortunately, it appears that this person's productivity was due in part to wholesale combining of questions and shortening of interviews with serious implications for data quality. The same variation in follow-up procedures for in-movers and WHUHEs was also found in the Alaska Native site and the Hispanic site.

The ethnographer reported no clear cases of possible errors, though a few were unresolved. Not one person in her sample of rostered persons was marked yes for either sometimes living or staying somewhere else (question 7) or for having a possible omission on April 1, 2010 (question H1). Most of her respondents were elderly and in small apartments.

## **5.2 The Census Coverage Measurement Person Interview Component**

In this section, results from the 2010 CCM Person Interview Component of this evaluation are presented. The CCM is an independent post-enumeration survey that is designed to estimate net census coverage as a whole and also for certain populations, such as race and Hispanic populations and those based on sex and age. This survey is independent of the 2010 Census and includes its own listing, data collection operations, and processing. The national sample size was less than half of one percent of all U.S. housing units. The CCM household and person records are matched to their counterparts in the 2010 Census dataset to estimate the accuracy of the 2010 Census. Because the CCM interview is more complex regarding coverage than the initial census interview,<sup>13</sup> observing and audiotaping (when permitted) this live interview and debriefing

---

<sup>13</sup> A brief description of the CCM Person Interview Operation and summary of the instrument is in Section 3.1.4.

afterwards provided a unique opportunity to examine the situations and identify factors that may lead to differential coverage of specific race/ethnic groups.

The section follows the same outline as the last, starting with a brief overview of the field observation phase in which ethnographers observed live interviews and debriefed respondents immediately after the interview and documenting the numbers of ethnographers and enumerators accompanied, as well as the numbers of interviews observed and audiotaped and showing variation across the race/ethnic sites. The focus then turns to the quantitative analysis of possible coverage error using pooled data across sites from separate data sources. By validating the final localized unedited records with field observations and debriefings, as well as with the independent in-depth CCM data already matched to the corresponding 2010 Census records,, it is possible to examine types and sources of possible coverage errors among records and characteristics of persons and households affected by it who may be at higher risk of coverage error. The focus then widens to identify cross-cutting themes from the insights in the ethnographers' site reports as well as profiles of factors in each race/ethnic site that shed light on why some groups may be consistently miscounted.

During mid- to late-August, 2010, the second set of nine ethnographers accompanied 53 interviewers to observe 318 interviews in the eight race/ethnic sites included in the 2010 Census observations in the last section, as well as a new Middle Eastern site in the Wayne and Macomb County areas of Michigan, including parts of Detroit and Dearborn. As shown in Table 14, the number of interviews observed per site ranged from 26 in the Hispanic site to 45 on the two American Indian reservations.

The numbers of interviews in each site that were conducted with a respondent from the target group are also shown in Table 14. They range from 6 (19 percent) of the 31 in the Middle Eastern site to 44 (98 percent) of 45 respondents in the American Indian site. Overall, 160 (57 percent) of 281 interviews in the eight sites with targeted race/ethnic groups were conducted with respondents of that target group. If the races of all persons in the households were considered, the Alaska Native number of households with Alaska Natives rises from 12 to 16, the American Indian number of households with American Indians would increase from 44 to 45, and the Native Hawaiian number of households with Native Hawaiians or Other Pacific Islanders would go up from 11 to 20, for a total of 178 households (63 percent) of 281 households in the same eight sites including one or more persons of the target race/ethnic group.

Of these 318 interviews, 252 were audiotaped and summarized, while the remaining interviews were summarized. A total of 247 (79 percent) of these were audiotaped. The percentages of all interviews that were audiotaped in each site is as follows: African American sites<sup>14</sup> (n = 23, 85 percent); Alaska Native site (n = 30, 86 percent) American Indian site (n = 38, 84 percent); Asian site, focused more on Chinese (n = 32, 74 percent); Middle Eastern site (n = 22, 71 percent); Native Hawaiian/Other Pacific Islander site (n = 29, 81 percent); Non-Hispanic White site (n = 31, 82 percent); Hispanic (n = 18, 69 percent); and Generalized site (n = 29, 78 percent). It should be noted that these are distributions by *geographical site*, not by *race/ethnicity*. The

---

<sup>14</sup> Observations of CCM interviews were conducted on two American Indian reservations: the reservation identified as Site 1 in the NRFU/UE section and another American Indian reservation. This change was necessary in order to observe at least 35 CCM interviews.

reader should keep in mind that each site included persons outside the target group, as shown in Table 14.

Table 14: CCM Interviewers Accompanied and Interviews Observed by Race/Ethnic Sites

| Target Race/ethnic Group in Site            | Location                           | Interviewers accompanied | Interviews observed |  |
|---|------------------------------------|--------------------------|---------------------|--|
|   |                                    |                          | Total               | Interviews with Respondents of Target Race/Ethnic Group for Site |
| African American                            | Chicago, Illinois                  | 7                        | 27                  | 24   |
| Alaska Native*                              | Kodiak Island Borough, Alaska      | 1                        | 35                  | 12   |
| American Indian*                            | Two Southwest reservations         | 5                        | 45                  | 45   |
| Asian (mostly Chinese)                      | San Francisco Bay Area, California | 8                        | 43                  | 17   |
| Middle Eastern                              | Detroit, Dearborn, Michigan        | 5                        | 31                  | 6  |
| Native Hawaiian and Other Pacific Islander* | Big Island, Hawaii                 | 7                        | 36                  | 11   |
| Non-Hispanic White                          | Jackson County, Missouri           | 9                        | 38                  | 36   |
| Hispanic                                    | Dallas/Fort Worth, Texas           | 6                        | 26                  | 10   |
| Generalized site                            | Broward County, Florida            | 5                        | 37                  | No target group  |
|   |                                    |                          |                     |  |
| <b>TOTAL</b>                                |                                    | <b>53</b>                | <b>318</b>          | <b>160</b>   |

Sources: Comparative Ethnographic Studies CCM PI Ethnographer Reports

\*Taking into account the race/ethnic identifications of all persons in the households, the Alaska Native site had 16 households with any part or full Alaska Natives and the Native Hawaiian site had 20 households with part or full Native Hawaiians and/or Other Pacific Islanders. All of the American Indian interviews had households with part or full American Indians.

### 5.2.1 Records check: numbers, types, and reasons for possible coverage errors

This section presents results on possible coverage error in the ethnographer-observed CCM PI interviews operation across all sites. After the operations were complete, Census Bureau staff matched the observed CCM PI households and persons to the final localized 2010 Census

records in the pre-matched 2010 CCM/Census datasets to do a partial validation study of the 2010 Census in the observed housing units. The purposes were to triangulate across data sources to identify consistencies and inconsistencies on who should be counted at the observed housing units and to attempt to identify possible sources of error, using information from all sources.

To determine if and where each person should be counted, our coding team compared rosters from four data sources: 1) the roster collected in the observed standard interview; 2) the roster of persons the ethnographer assessed should be counted there and elsewhere; 3) the roster included in the final 2010 Census unedited file that was pre-matched to the 4) the final Coverage Measurement assessment of where each person should be counted. Possible coverage errors were identified by inconsistencies in the records of where a person was counted and where a person should have been counted according to the team who reviewed all sources. Difficult cases were discussed by the team and the authors reviewed some coding. If not enough information was available, a person was coded as undetermined.

In addition to identifying where a person should have been counted, coders identified possible sources of coverage error, given all record sources. We identified persons and households as having possible coverage error by identifying inconsistencies between data sources. We offer the following limitation. In this study, all categories of the coverage error variable indicate *possible* coverage errors. We use caution when making statements about coverage error and note the possibility that neither the team assessment nor the ethnographer may be the most knowledgeable about where a person should be counted. Some persons' records may include updates from follow-up interviews after the ethnographer-observed interviews, which may result in a more informed determination for where the person should be counted.

Our analysis sample consists of persons who: 1) likely should have been counted at the observed housing unit on Census Day and appeared to have been correctly counted there in the final localized 2010 Census dataset matched to the CCM; 2) likely should have been counted at the observed housing unit on Census Day but appeared to have been omitted; or 3) likely should not have been counted at the observed housing unit on Census Day but appeared to have been incorrectly counted there.<sup>15</sup>

The rest of this section will focus on the analysis sample. At the household level, the analysis sample consisted of 288 households.<sup>16</sup> Regarding household tenure on the day of the CCM interview, respondents most-often reported housing units as owned with a mortgage or loan (44 percent), or rented (28 percent). In addition, the most-frequent household sizes were five or more persons (31 percent) and two persons (24 percent). Comparing across race/ethnic sites, the largest households were located on the second American Indian Reservation, which had a median

---

<sup>15</sup> The authors omitted 213 persons identified in observed housing units who should have been counted elsewhere and appear to have been correctly not counted at the observed housing unit for two reasons: (1) a number of them had moved outside the geographical area covered by our CCM records check dataset and this precluded our assessing these cases consistently and (2) the ethnographers did not observe or debrief respondents in those other units. All of these types of persons are included in the official national CCM estimates.

<sup>16</sup> The authors excluded 30 observed units recorded as having no population on Census Day: those classified as vacant, not a housing unit, and whole household usual home elsewhere.

household size of six persons. Included in analysis households are also two group quarters that were enumerated as housing units, and one household did not have enough information to determine household type on April 1, 2010.

The overall analysis sample included 953 persons in the 288 households with the following race and ethnic distributions (shown in Table 15). Of these, 11 percent were recorded in the CCM dataset as Hispanics of any race. Persons were distributed across non-Hispanic race groups as follows: Whites comprised 29 percent, American Indian and Alaska Natives were percent, Blacks were 15 percent, Asians were 11 percent, Other was 2 percent, and Native Hawaiian and Other Pacific Islander were 2 percent. A larger-than-expected 6 percent reported of non-Hispanic multiple races, almost all of whom were in the Native Hawaiian site and the Alaska Native site. The remaining 8 percent had missing data for race and ethnicity.

It is important to keep in mind that each of the targeted sites had a mix of persons of different races and/or Hispanic origin. Some sites, such as the American Indian sites, the non-Hispanic White site and African American site had a very large majority of respondents in the target race/ethnic group, compared to others that had relatively low proportions, such as the Asian site and the Alaska Native site. It is therefore not appropriate to equate the name of the targeted race/ethnic *site* with the incidence of coverage error for that race/ethnic *group*.

The most frequent age cohorts represented were between ages 30 and 49 years old (26 percent) and persons aged 50 and older (24 percent). Most persons (89 percent) were added to the roster using the initial roster question. The next most-frequent roster probe that added persons to the standard interview roster was the final “Anyone else?” probe in the series (3 percent). The most-frequent types of resident were continuous residents (87 percent) and outmovers (6 percent).

Regarding alternative addresses, most persons had no other alternative address (76 percent). Just over one-tenth of persons had one alternative address (13 percent), which were most frequent at the Native Hawaiian site.



Table 15: Analysis Sample Distribution of Race and Ethnicity across CCM Race/Ethnic Sites

| Race/ Ethnicity Category                               | Race/Ethnic Site |           |            |            |           |            |           |           |            |           |            |
|--|------------------|-----------|------------|------------|-----------|------------|-----------|-----------|------------|-----------|------------|
|  | AA               | AI1       | AI2        | AN         | AS        | GS         | HI        | ME        | NH         | WH        | Total      |
| Non-Hispanic White                                     | 0                | 0         | 0          | 30         | 27        | 27         | 18        | 49        | 27         | 77        | <b>255</b> |
| Non-Hispanic Black                                     | 53               | 0         | 0          | 0          | 1         | 30         | 10        | 33        | 0          | 0         | <b>132</b> |
| Non-Hispanic American Indian/Alaska Native             | 0                | 53        | 140        | 11         | 2         | 0          | 0         | 0         | 0          | 3         | <b>209</b> |
| Non-Hispanic Asian                                     | 0                | 0         | 0          | 35         | 27        | 2          | 4         | 0         | 27         | 1         | <b>96</b>  |
| Non-Hispanic Native Hawaiian or Other Pacific Islander | 0                | 0         | 0          | 0          | 0         | 0          | 0         | 0         | 16         | 0         | <b>16</b>  |
| Non-Hispanic Other                                     | 0                | 0         | 0          | 0          | 2         | 2          | 1         | 7         | 5          | 0         | <b>17</b>  |
| Non-Hispanic Multiple Race                             | 0                | 0         | 0          | 18         | 0         | 1          | 1         | 0         | 28         | 1         | <b>49</b>  |
| Hispanic White   | 0                | 53        | 1          | 6          | 6         | 25         | 4         | 0         | 4          | 0         | <b>46</b>  |
| Hispanic Black   | 0                | 0         | 0          | 0          | 0         | 3          | 0         | 0         | 0          | 0         | <b>3</b>   |
| Hispanic American Indian/Alaska Native                 | 0                | 0         | 6          | 0          | 0         | 0          | 5         | 0         | 0          | 0         | <b>11</b>  |
| Hispanic Asian   | 0                | 0         | 0          | 0          | 0         | 0          | 0         | 0         | 1          | 0         | <b>1</b>   |
| Hispanic Native Hawaiian/Other Pacific Islander        | 0                | 0         | 0          | 0          | 0         | 0          | 0         | 0         | 4          | 0         | <b>4</b>   |
| Hispanic Other   | 6                | 0         | 1          | 3          | 6         | 11         | 3         | 0         | 2          | 0         | <b>32</b>  |
| Hispanic Multiple Race                                 | 0                | 0         | 0          | 1          | 0         | 0          | 3         | 0         | 3          | 0         | <b>7</b>   |
| <b>Total</b>   | <b>59</b>        | <b>53</b> | <b>148</b> | <b>104</b> | <b>71</b> | <b>101</b> | <b>49</b> | <b>89</b> | <b>117</b> | <b>87</b> | <b>878</b> |

Note: 75 persons have missing data for combined Hispanic origin and race. Source: Comparative Ethnographic Studies CCM PI Analysis Sample Dataset. AA = African American site, AI = American Indian site, AN = Alaska Native site, AS = Asian site, GS site = Generalized site, HI = Hispanic site, ME = Middle Eastern site, NH = Native Hawaiian and Other Pacific Islander site, WH = non-Hispanic White site.

### 5.2.1.1 Numbers and Types of Possible Coverage Errors in the Records Check

Table 16 shows that 873 persons (94 percent) of the 953 analysis sample persons did not have an indication of any type of coverage error and appeared correctly enumerated at the observed housing unit. Of the remaining persons, 60 (6 percent) had an indication of some type of possible coverage error, including people who were likely omitted or counted in the wrong place. Finally, 20 other persons were coded as “Unresolved” or “Missing” because there was not enough information to in the final localized 2010 CCM/Census matched dataset to determine where they were counted by the 2010 Census, likely should have been counted according to the team assessment, or both.

Table 16 also shows the types of possible coverage error identified through the records check when the team assessment and CUF variables differed in terms of whether a person should be counted at the observed housing unit: 32 had possible omissions at the OHU, 24 had possible incorrect counts, four were possible overcounts, meaning that they were out of scope for the 2010 Census (e.g., born/immigrated after Census Day, died/emigrated before Census Day), 11 were unresolved and nine had missing values.

Table 16: Observed CCM Persons with Possible Coverage Error

| Possible Coverage Error Type*                        | Persons    |            |
|--|------------|------------|
|  | Frequency  | Percent    |
| No Coverage Error at OHU in localized dataset        | 873        | 92         |
| Possible Incorrect Count at OHU in localized dataset | 24         | 3          |
| Possible Overcount at OHU in localized dataset       | 4          | < 1        |
| Possible Omission at OHU in localized dataset        | 32         | 3          |
| Unresolved   | 11         | 1          |
| Missing Data   | 9          | 1          |
|  |            |            |
| <b>Total</b>   | <b>953</b> | <b>100</b> |

Note: OHU = observed housing unit. Percentages are rounded to the nearest whole number.

Source: Comparative Ethnographic Studies CCM PI Analysis Sample Dataset

\* Possible coverage error type refers to possible coverage error as found in the final localized 2010 CCM/Census matched dataset. It is possible that some number of persons with possible coverage error were correctly counted in the 2010 Census elsewhere in the country.

#### 5.2.1.1.1 Possible omission at the observed housing unit

Thirty-two persons in 20 households had a possible omission at the observed housing unit; a person should have been counted there, but records indicate that the person was counted someplace else or not at all. A separate analysis of the ethnographer transcripts and summaries shows two main categories of possible omissions.

The first of the main omission categories included 24 persons identified in the CCM standard interview or in the final localized 2010 CCM/Census matched dataset who could not be linked to persons in the matched 2010 Census dataset for those observed households. This group yielded 24 persons in 15 housing units. In five of these housing units, the whole household appeared to be missed, accounting for 11 persons. In the remaining 10 housing units, some but not all persons were possible omissions, totaling 13 persons.

In two cases, one in the Middle Eastern site and the other in the Alaska Native site, the CCM respondents included themselves in the observed households on Census Day, but they were not found in the localized matched 2010 Census records for the same households in the final localized 2010 CCM/Census matched dataset, which were reported by other persons in the households. In other cases, all of the evidence in the CCM answer patterns indicated that persons should have been counted there on Census Day, providing no cues as to why persons might have been omitted from the final 2010 Census rosters in the final localized 2010 CCM/Census matched dataset. Some of these possible omissions were for everyone in the household, while others were just for some of the persons. For example, in one Native Hawaiian site case, both the husband and wife were on the CCM roster but not counted on the final 2010 Census roster. In another in the Native Hawaiian site, eight persons were listed in CCM, but only five were included in the 2010 Census final dataset in our localized dataset: the uncle and two of the respondent's young biological children, one of them less than a month old, were possible omissions in the 2010 Census. Similarly, on one American Indian reservation, a baby of less than 2 months identified in the CCM was not included on the final matched 2010 Census roster, though his mother was.

The second group of possible omissions consisted of six persons in three OHU who were identified by the ethnographer in a debriefing immediately after the CCM interviewer completed the standard interview. Because these are persons who were not captured in the CCM standard instrument, these persons do not show up in the final overall CCM sample. These include four outmovers missed in one interview in the Non-Hispanic White site due to interviewer error with the instrument, one person missed in the Alaska Native site in a hidden household that had not been included in either the 2010 Census or the CCM operations, and one possible outmover in the Native Hawaiian site who remained unnamed in the standard interview, but who was identified by the ethnographer in the debriefing. We did not find these persons in these or alternative 2010 Census households in the final localized 2010 CCM/Census matched dataset. Had the interviewer probed the respondent a bit more, she could have gotten that name. It is not clear if that former owner was actually still there on April 1, 2010. This possible outmover had not been listed here during the NRFU operation, apparently because a NRFU enumerator mistakenly counted the later in-movers as already living in the OHU on April 1, 2010 when they had not yet moved in and the outmover may still have been there. Two other persons in two households were possible omissions.

#### 5.2.1.1.2 Possible incorrect counts at the observed housing unit

As noted earlier, 24 persons were coded as possible incorrect counts at the observed housing unit per the 2010 Census records in the final localized 2010 CCM/Census matched dataset, but the person should be counted at another place. They are associated with 17 housing units. In three of

the 17 housing units, the whole household appeared to be incorrectly counted at the OHU, for a total of six persons. In the remaining 14 households, just some of the 18 persons appeared to be incorrectly counted here.

This category includes the three persons previously identified in the Native Hawaiian site who were counted in NRFU as having lived in the OHU on Census Day, but who did not move in until mid-April; they should have been counted at their former place. In the Alaska Native site, a female boarder who had moved out to some undisclosed location near the end of March was still included on the OHU mailout form and apparently miscounted there in the 2010 Census. The respondent stated she and the boarder had completed and submitted the 2010 Census mailout form in mid-March, when the boarder was still living there. Had she waited until April 1, 2010 to complete the form, she would not have rostered the boarder. The coder/matchers did not find any additional address for this boarder, so it appears she was included in the 2010 Census, but in the wrong place.

This interview shows the problem of possible errors that can occur in the 2010 Census when the Census Bureau sends out the mailout forms in mid-March and continues to prompt respondents to mail the form in as soon as possible. Had this respondent waited until April 1, 2010 to send her form in, she would have likely completed her 2010 Census form correctly without this woman. However, there is a chance that this woman could then have been missed in the 2010 Census; the coding/matching team member searched the PerMaRCS dataset for alternative addresses but did not find this woman counted at another place.

This same Alaska Native case also illustrates the point that mailout 2010 Census forms submitted before April 1, 2010 may include persons who were either erroneously included on or omitted from the roster of who would actually be living or staying there on Census Day. A graph of cumulative mail return rates in the 2010 Census indicates that the Census Bureau had received mailout forms from about 60 percent of occupied housing units on or before Census Day, April 1, 2010 (Letourneau 2012: Appendix I, page 58). Some unknown portion of mailout forms may have possible errors relative to where persons should have been counted on April 1, 2010.

In another case in the non-Hispanic White site, a woman started staying overnight at her daughter's home for caregiving around March 1, 2010 when her daughter was diagnosed with cancer. While she had not planned to do so, she stayed there continuously and was still there on the CCM Interview Day of September 1. The very interesting situation here is that the CCM instrument asks for the amount of time spent in the alternative address over the year backward from the CCM interview date, which would make the CCM reference period September, 2009 to August 31, 2010. This woman stayed just over six months of this time period at the daughter's place, so the instrument would likely indicate she should be counted at her daughter's place. However, she clearly said that neither she nor anyone at her daughter's place would consider her to "live" there; she "stayed" there. This highlights a problem with the CCM trying to determine a person's usual residence for the 2010 Census with a reference period that extends beyond April 1, 2010. From the 2010 Census residence rule standpoint, it is the status of the person on April 1, 2010 that is the important factor. Any residence changes after that date are irrelevant in determining where a person should be counted in the 2010 Census.

### 5.2.1.2 Reasons for Possible Coverage Errors in Records Check

Coder/matcher team members coded each person for sources of possible coverage error. Sources of possible coverage error are reasons why possible coverage errors of persons may have happened. Up to three sources of coverage error were coded for each person. Of 60 persons who had an indication of some type of coverage error, team members coded 24 different sources of coverage error a total of 91 times. The most-frequent sources of coverage error were respondent confusion (n = 16) and inconsistent living arrangements (n = 13). See Table 17 for sources of coverage error for persons with possible coverage error.

Sources of possible coverage error for persons were also collapsed into eight broad categories, as shown in Table 18. The category, “mobility and tenuousness,” includes people who moved, people with more than one place to live and people who were tenuously attached to the household. At the collapsed level, “mobility/tenuousness” was the largest source of possible error, and was often reported in connection with respondent confusion. Respondent confusion was coded when the respondent provided inconsistent responses or was observed answering a question different from the one that was asked; this could have led to a coverage error. For example, if an interviewer asked “Did you move here before or after April 1, 2010?” and the respondent answered, “After...we moved here about March 15<sup>th</sup>.” – this would have been coded as “respondent confusion.” Interviewer error included omitting or rewording questions, making assumptions (answering without waiting for the respondent to answer), not probing sufficiently, not otherwise adhering to the interview protocol, and sometimes not following up with a respondent. Respondent concealment or refusal was identified when respondents refused to answer questions or otherwise indicated they were being untruthful when answering coverage-related questions. Hidden housing units were those identified in the CCM or by the ethnographer that had not been identified through previous 2010 Census operations. Language barriers were identified when a lack of speaking the same language resulted in a possible coverage error.

Table 17: Sources of Possible Coverage Error for Persons with Possible Coverage Error in this Evaluation

| Source of Possible Coverage Error  | Frequency | Percent    |
|--|-----------|------------|
| Respondent confusion   | 16        | 18         |
| Inconsistent living arrangements   | 13        | 14         |
| Ethnographer did not probe sufficiently  | 7         | 8          |
| Respondent concealment   | 6         | 7          |
| Person not included on Census form but respondent mentioned person during the CCM interview                                | 5         | 6          |
| Inmover status   | 5         | 6          |
| Outmover status  | 5         | 6          |
| Unknowledgeable proxy  | 5         | 6          |
| Omission of question   | 3         | 3          |
| Foreign language translation of interview (other than Spanish)   | 3         | 3          |
| Address missed during previous 2010 Census operations  | 3         | 3          |
| Respondent refusal to answer   | 3         | 3          |
| Question rewording   | 2         | 2          |
| Leading Question   | 2         | 2          |
| Interviewer error: what is recorded in the PerMaRCS CCM record does not match what was said in the CCM interview.          | 2         | 2          |
| Instrument error: the interviewer appeared to ask all the correct questions but the data appear to be correct in PerMaRCS. | 2         | 2          |
| Moving close to but before Census Day  | 2         | 2          |
| Poorly worded CCM question   | 1         | 1          |
| Having two houses  | 1         | 1          |
| Individual temporarily living in a GQ on or around Census Day  | 1         | 1          |
| Cycling for custody  | 1         | 1          |
| Interviewer did not probe sufficiently   | 1         | 1          |
| Baby or small child around Census Day  | 1         | 1          |
| Moving on unspecified date around Census Day   | 1         | 1          |
|  |           |            |
| <b>Total</b>   | <b>91</b> | <b>100</b> |

Notes. Team members did not code sources for 17 persons. Percentages are rounded to the nearest whole number. Note: AA = African American site, AI = American Indian site, AN = Alaska Native site, AS = Asian site, GS = Generalized site, HI = Hispanic site, NH = Native Hawaiian and Other Pacific Islander site, and WH = Non-Hispanic White site.

Source: Comparative Ethnographic Studies CCM PI Analysis Sample Dataset

Five cases showed no evidence of why a possible coverage error should have occurred (other than that the records between operations were inconsistent), and 16 cases showed other reasons

for possible coverage errors, including omitting babies, computer glitches, and respondent frustration.

Table 18: All Collapsed Team Sources of Possible Coverage Error for Persons with Possible Coverage Error in Observed Sites

| Source of Possible Coverage Error | Frequency | Percent    |
|-----------------------------------|-----------|------------|
|                                   |           |            |
| Mobility/Tenuousness              | 29        | 32         |
| Respondent confusion              | 16        | 18         |
| Interviewer error                 | 10        | 11         |
| Respondent concealment/refusal    | 9         | 10         |
| Hidden housing unit               | 3         | 3          |
| Language barrier                  | 3         | 3          |
|                                   |           |            |
| Unknown                           | 5         | 5          |
| Other                             | 16        | 18         |
|                                   |           |            |
| <b>Total</b>                      | <b>91</b> | <b>100</b> |
|                                   |           |            |

Note: Percentages are rounded to the nearest whole number.

Source: Comparative Ethnographic Studies CCM PI Analysis Sample Dataset

In order to understand more fully how these sources of possible error contributed to coverage error, possible coverage error types were cross tabulated by source of possible coverage error. Table 19 shows sources of error for possible omissions. Table 20 shows the combined sources of error for people coded as possibly erroneously counted at the observed housing unit when they should have been counted elsewhere or not included in the 2010 Census.

Table 19: Team Sources of Possible Omissions for Persons with Possible Coverage Error in Observed Sites

| Source of Possible Coverage Error | Frequency | Percent    |
|-----------------------------------|-----------|------------|
|                                   |           |            |
| Mobility/Tenuousness              | 12        | 29         |
| Respondent concealment/refusal    | 4         | 10         |
| Respondent confusion              | 3         | 7          |
| Hidden housing unit               | 3         | 7          |
| Language barrier                  | 3         | 7          |
| Interviewer error                 | 2         | 6          |
| Unknown                           | 5         | 12         |
| Other                             | 9         | 22         |
|                                   |           |            |
| <b>Total</b>                      | <b>41</b> | <b>100</b> |

Note: Percentages are rounded to the nearest whole number.

Source: Comparative Ethnographic Studies CCM PI Analysis Sample Dataset

Table 20: Team Sources of Possible Overcounts/Incorrect Counts for Persons with Possible Coverage Error in Observed Sites

| Source of Possible Coverage Error | Frequency | Percent    |
|-----------------------------------|-----------|------------|
|                                   |           |            |
| Mobility/Tenuousness              | 17        | 34         |
| Respondent confusion              | 13        | 26         |
| Interviewer error                 | 8         | 16         |
| Respondent concealment/refusal    | 5         | 10         |
| Other                             | 7         | 14         |
|                                   |           |            |
| <b>Total</b>                      | <b>50</b> | <b>100</b> |

Note: Percentages are rounded to the nearest whole number.

Source: Comparative Ethnographic Studies CCM PI Analysis Sample Dataset

In this small study, all instances of hidden housing units and language barriers are linked to possible omissions. Situations involving mobility or tenuousness, respondent confusion, concealment and interviewer error are linked to both omissions and miscounts.

De la Puente (1993; 2003) cited language barriers and households with mobility, as well as irregular housing (including hidden housing units) as reasons for undercoverage in both the 1990 Census and Census 2000 studies. Respondent confusion and respondent concealment have been cited as reasons for undercoverage since the earliest ethnographic studies (Aschenbrenner, 1975; Valentine and Valentine, 1971). By having ethnographers actually observe the interviews in the



current study, interviewer error was able to be identified as a direct contributor to both types of possible coverage error - omissions and miscounting.

## **5.2.2 Characteristics of persons and households with possible coverage errors**

### **5.2.2.1 Persons with Possible Coverage Errors**

Most persons with possible coverage error were brought on to the roster with the initial roster question (n = 44), and had one alternative address (n = 28). In terms of relationship type to the reference person, the reference person's biological son or daughter (n = 14, 6 percent) had the most instances of possible coverage error. The aunt/uncle (n = 1, 20 percent), roomer/boarder (n = 1, 20 percent), housemate/roommate (n = 2, 18 percent), and children for whom relationship to Person 1 could not be determined (n = 1, 17 percent) were the relationship types that had the greatest proportions of possible coverage error.

Across all sites, the 0-4 age cohort (n = 7, 15 percent) and the 18-29 age cohort (n = 13, 12 percent) had the greatest proportions of persons within their own cohorts with possible coverage error. A chi-square test of independence also found that instances of possible coverage error were associated with age cohort ( $X^2 = 22.93$ ,  $DF = 6$ ,  $p = .0008$ ). These age ranges have been associated with higher levels of coverage error in past decennial studies. The overall CCM coverage results show a significant undercount of children from ages 0 to 4 and young adult males from ages 18 to 29 of 0.72 percent and 1.21 percent respectively, while young women from ages 18 to 29 showed a slight, but non-significant overcount of 0.28 percent (U.S. Census Bureau May 22, 2012 Press Release, slides 21, 22).

### **5.2.2.2 Housing Units and Households with Possible Coverage Error**

As mentioned, 41 households had at least one person with possible coverage error. Half of these 41 households had five persons or more, even though these large households comprised just 31 percent of the sample. Households with five persons or more were significantly more likely to have possible coverage error when those other households are placed in a single group ( $X^2 = 9.62$ ,  $DF = 1$ ,  $p < .01$ ). These findings are consistent with those from past censuses in finding large households to be at greater risk of coverage error.

Past research has also documented that coverage errors are more likely to occur in complex than noncomplex households (e.g., de la Puente, 1993; Schwede, Blumberg, and Chan, 2006). Households were classified into the complex/noncomplex household typology whereby a nuclear family consisting of married parents and their joint biological children, a single parent and his/her biological children, and a person living alone were considered noncomplex and all others as complex (Schwede, 2008b). Multi-generational households constitute one complex household category that varies across race/ethnic groups. According to 2010 Census results, four percent of U.S. households were multigenerational, ranging from three percent for non-Hispanic Whites alone to 11 percent for Native Hawaiians/Other Pacific Islanders and for Some Other Race (Schwede, 2012).

In the overall analysis sample of 288 analysis sample households, 105 were complex and 180 were noncomplex.<sup>17</sup>

The second American Indian Reservation (n = 19) and the Native Hawaiian site (n = 16) had the largest numbers of complex households. That same second American Indian Reservation and the first American Indian Reservation had the highest proportions of all households that were complex – 68 percent and 64 percent, respectively, while the Native Hawaiian site had 44 percent. See Table 21 for a distribution of complex and noncomplex households across CCM race/ethnic sites. The Generalized site had the lowest proportion, with six (17 percent) of 35 households complex

Table 21: Census Day Household Type for all Analysis Sample CCM households by Race/Ethnic Sites in Evaluation

| <b>Race/Ethnic Site</b>  | <b>Complex Households</b> | <b>Noncomplex Households</b> | <b>Total</b> |
|--------------------------|---------------------------|------------------------------|--------------|
| American Indian (Site 2) | 19                        | 9                            | <b>28</b>    |
| Native Hawaiian          | 16                        | 20                           | <b>36</b>    |
| Alaska Native            | 13                        | 18                           | <b>31</b>    |
| Asian                    | 13                        | 20                           | <b>33</b>    |
| Middle Eastern           | 11                        | 16                           | <b>27</b>    |
| American Indian (Site 1) | 9                         | 5                            | <b>14</b>    |
| Hispanic                 | 6                         | 17                           | <b>23</b>    |
| African-American         | 6                         | 16                           | <b>22</b>    |
| Generalized Site         | 6                         | 29                           | <b>35</b>    |
| White                    | 6                         | 30                           | <b>36</b>    |
| <b>Total</b>             | <b>105</b>                | <b>180</b>                   | <b>285</b>   |

Note: Two households were excluded because they were group quarters. One household was excluded because of missing data for household type.

Source: Comparative Ethnographic Studies CCM PI Analysis Sample Dataset

Of 41 households with any possible coverage errors, 18 were noncomplex and 22 were complex.<sup>18</sup>

The most-frequent households were at the Native Hawaiian site (n = 11) and the Alaska Native site (n = 7). The American Indian sites, the Alaska Native site, the Middle Eastern site, and Native Hawaiian site had more complex households than noncomplex households with coverage error, while the opposite was the case in the non-Hispanic White site, the Generalized site, and the Asian site. In the African American site and the Hispanic site, there were the same numbers of complex households and noncomplex households with possible coverage error. See Table 22 for the household type of all CCM households with possible coverage error across race/ethnic sites.

<sup>17</sup> Three units lacked information to classify by household complexity.

<sup>18</sup> One unit lacked information to classify by household complexity.

Table 22: Census Day Household Type for CCM Households with Possible Coverage Error across Race/Ethnic Sites in this Evaluation

| <b>Race/Ethnic Site</b>  | <b>Complex Households</b> | <b>Noncomplex Households</b> | <b>Total</b> |
|--------------------------|---------------------------|------------------------------|--------------|
| African-American         | 2                         | 2                            | <b>4</b>     |
| American Indian (site 1) | 2                         | 0                            | <b>2</b>     |
| American Indian (site 2) | 4                         | 0                            | <b>4</b>     |
| Alaska Native            | 5                         | 2                            | <b>7</b>     |
| Asian                    | 0                         | 1                            | <b>1</b>     |
| Generalized              | 0                         | 3                            | <b>3</b>     |
| Hispanic                 | 1                         | 1                            | <b>2</b>     |
| Middle Eastern           | 1                         | 0                            | <b>1</b>     |
| Native Hawaiian          | 6                         | 5                            | <b>11</b>    |
| Non-Hispanic White       | 1                         | 4                            | <b>5</b>     |
| <b>Total</b>             | <b>22</b>                 | <b>18</b>                    | <b>40</b>    |

Note: One household was excluded because it was a group quarters.

Source: Comparative Ethnographic Studies CCM PI Analysis Sample Dataset

To determine whether possible coverage error differed by household type on April 1, 2010, a chi-square test of independence was conducted to investigate whether possible coverage error was associated with household type. It was hypothesized that possible coverage error would be associated with household type, and that possible coverage error would be found more often in complex households than in noncomplex households. The chi-square test found that possible coverage error was associated with household type ( $X^2 = 6.59$ ,  $DF = 1$ ,  $p < .01$ ), and that there were more instances of possible coverage error in complex households than in noncomplex households (i.e., 22 instances versus 18 instances, respectively).

### 5.2.3 Distribution of possible coverage errors across race/ethnic groups and sites

Of the 953 persons observed, 60 were judged to exhibit some form of possible coverage error, while 11 others were unresolved. The race/ethnic groups with the most persons with possible coverage error were Whites ( $n = 16$ ), Hispanics of any race ( $n = 5$ ), Blacks ( $n = 7$ ), American Indian/Alaska Natives ( $n = 7$ ), and Native Hawaiians and Other Pacific Islanders ( $n = 5$ ), as shown in Table 23. To test significance, individual chi-square tests of independence were conducted to compare the proportion of people in a race/ethnic group with possible coverage error with all other persons placed in one group. In these analyses, Hispanics of all races were in the Hispanic group, and all race groups included persons who were Hispanic and non-Hispanic of that race. As shown in Table 23, the only race/ethnic group significantly more likely to have coverage error than the others was Native Hawaiian/Other Pacific Islander, of whom 25 percent had possible coverage error ( $X^2 = 16.02$ ,  $DF = 1$ ,  $p < .01$ ). Coverage error levels for the other

race and ethnic groups were not significantly different from all others. These results should not be interpreted to suggest there is something inherent in being Native Hawaiians/Other Pacific Islanders that causes coverage errors; rather, factors are presumed to be associated with living situations more often observed among Native Hawaiians in this site that were associated with possible coverage errors.

Table 23: Frequency of Persons with Possible Coverage Error by Race/Ethnic Group in this Evaluation

| <b>Race/Ethnic Group</b>               | <b>Total in Race/Ethnic Group</b> | <b>Persons With Possible Coverage Error</b> | <b>Percent</b> |
|--|-----------------------------------|---|----------------|
| Native Hawaiian/Other Pacific Islander | 20                                | 5   | 25*            |
| Other Race/Ethnic Group                | 42                                | 4   | 10             |
| Hispanics of any race                  | 110                               | 8   | 7              |
| Multiple Races                         | 59                                | 4   | 7              |
| White                                  | 284                               | 15  | 5              |
| Black                                  | 134                               | 7   | 5              |
| Middle Eastern                         | 26                                | 1   | 4              |
| American Indian/Alaska Native          | 213                               | 7   | 3              |
| Asian                                  | 97                                | 3   | 3              |
|  |                                   |   |                |

Note: \* =  $p < .01$ . Seventy-eight persons had incomplete race and Hispanic origin data. Of these persons, 14 had possible coverage error.

Source: Comparative Ethnographic Studies CCM PI Analysis Sample Dataset

The distribution of persons with possible coverage error of any type by site is shown in Table 24. They were most frequent at the Native Hawaiian site ( $n = 22$ ). Persons found to have been incorrectly counted at the observed housing unit were spread relatively evenly across sites, with seven of ten sites having at least two persons with this type of possible coverage error.

Most persons with this type of possible error were also at the Native Hawaiian site ( $n = 7$ ). Furthermore, persons with possible coverage error were found most at the Native Hawaiian site ( $n = 22$ ). A chi-square test of independence also found an overall association between race/ethnic site and instances of possible coverage error ( $X^2 = 35.59$ ,  $DF = 9$ ,  $p = <.0001$ ). In particular, the Native Hawaiian site was significantly associated with possible coverage error when compared to persons from all other sites ( $X^2 = 19.78$ ,  $DF = 1$ ,  $p = <.0001$ ).

The frequency of sources of possible coverage error by sites is shown in Table 25. The Native Hawaiian site had the highest frequency of sources of possible coverage error ( $n = 21$ ). The sites with the lowest frequencies were the Hispanic site, the Asian site, and the Middle Eastern site.

Table 24: Specific Types of Coverage Error for Persons across Race/Ethnic Sites in this Evaluation

| Specific Type of Coverage Error              | Race/Ethnic Site |          |          |           |          |          |          |          |           |           |           |
|--|------------------|----------|----------|-----------|----------|----------|----------|----------|-----------|-----------|-----------|
|  | AA               | AI1      | AI2      | AN        | AS       | GS       | HI       | ME       | NH        | WH        | Total     |
| Incorrectly Counted at Observed Housing Unit | 1                | 3        | 3        | 5         | 0        | 2        | 1        | 0        | 7         | 2         | <b>24</b> |
| Overcount at Observed Housing Unit           | 0                | 0        | 0        | 3         | 1        | 0        | 0        | 0        | 0         | 0         | <b>4</b>  |
| Possible Omission at Observed Housing Unit   | 5                | 1        | 1        | 2         | 0        | 1        | 1        | 1        | 13        | 7         | <b>32</b> |
| Unresolved                                   | 0                | 0        | 4        | 0         | 0        | 0        | 3        | 0        | 2         | 2         | <b>11</b> |
| <b>Total</b>                                 | <b>6</b>         | <b>4</b> | <b>8</b> | <b>10</b> | <b>1</b> | <b>3</b> | <b>5</b> | <b>1</b> | <b>22</b> | <b>11</b> | <b>71</b> |

Note: Nine persons have missing data for specific type of possible coverage error.

Source: Comparative Ethnographic Studies CCM PI Analysis Sample Dataset

In this section, the authors described the partial validation study with CCM data that triangulated among four data sources to decide where each person should be counted in the 2010 Census and compared that to where the person was counted in the final 2010 Census dataset matched to the CCM in the final localized 2010 CCM/Census matched dataset. The authors focused on the outcomes for the observed housing units on Census Day—those who were correctly or apparently incorrectly counted at the observed housing units or omitted from it when it appeared they should have been counted there. Others who were not counted at the OHU and should have been counted elsewhere were not included in the analysis. The results indicated that 60 people (6 percent) of the 953 people in the analysis sample had possible coverage error at the observed housing unit in the 2010 Census as of April 1, 2010. Thirty-two of them appeared to have been omitted from the observed housing unit while 24 appeared to have been incorrectly counted there and four were possible overcounts, meaning that they were out of scope for the 2010 Census. An additional 11 households were unresolved and 11 had missing data. At the household level, these 60 persons were found in 41 of the 289 housing units that were occupied on April 1, 2010. Some of these possible coverage errors were for the whole household while others were for just some household persons.

Table 25: Team Sources of Possible Coverage Error for Persons by Site in CCM Evaluation

| Coverage Error Source  | Race/Ethnic Site |           |          |           |          |          |          |          |           |           | Total     |
|--|------------------|-----------|----------|-----------|----------|----------|----------|----------|-----------|-----------|-----------|
|  | AA               | AI1       | AI2      | AN        | AS       | GS       | HI       | ME       | NH        | WH        |           |
| Respondent confusion   | 1                | 4         | 1        | 7         |          | 1        | 1        |          | 1         |           | 16        |
| Inconsistent living arrangement  | 3                | 4         | 1        | 3         |          | 1        |          |          |           | 1         | 13        |
| Ethnographer did not probe enough  |                  | 4         | 1        | 1         |          | 1        |          |          |           |           | 7         |
| Respondent concealment   | 2                |           |          | 2         |          |          |          |          | 2         |           | 6         |
| Person mentioned in CCM interview not listed on Census form  |                  |           |          |           |          |          |          | 1        | 4         |           | 5         |
| Inmover status   |                  |           |          |           |          |          |          |          | 5         |           | 5         |
| Outmover status  |                  |           |          |           |          |          | 1        |          |           | 4         | 5         |
| Unknowledgeable proxy  |                  |           |          |           |          |          |          |          |           | 5         | 5         |
| Omission of question   |                  |           | 1        | 2         |          |          |          |          |           |           | 3         |
| Foreign language translation (other than Spanish)  |                  |           |          |           |          |          |          |          | 3         |           | 3         |
| Address missed during previous 2010 Census operations  | 3                |           |          |           |          |          |          |          |           |           | 3         |
| Respondent refused   |                  |           |          | 2         | 1        |          |          |          |           |           | 3         |
| Question rewording   |                  |           |          | 2         |          | 1        |          |          |           |           | 2         |
| Leading Question   |                  |           | 1        |           |          |          |          |          |           |           | 2         |
| Interviewer error: PerMaRCS record does not match what was said in CCM interview                           |                  |           |          |           |          |          |          |          | 2         |           | 2         |
| Instrument error: interviewer seemed to ask all questions correctly but PerMaRCS data appear to be correct |                  |           |          |           |          |          | 1        |          | 1         |           | 2         |
| Moving just before April 1   |                  |           |          |           |          |          |          |          |           | 2         | 2         |
| Poorly worded CCM question   |                  |           |          |           |          | 1        |          |          |           |           | 1         |
| Having two houses  |                  |           |          |           |          |          |          |          | 1         |           | 1         |
| Person temporarily in a GQ on or around Census day   |                  |           |          |           |          |          |          |          |           | 1         | 1         |
| Cycling for custody  |                  |           |          |           |          |          |          |          | 1         |           | 1         |
| Interviewer did not probe enough   |                  |           | 1        |           |          |          |          |          |           |           | 1         |
| Baby or small child during Census Day  |                  |           | 1        |           |          |          |          |          |           |           | 1         |
| Moving on unspecified date around Census Day   |                  |           |          |           |          |          |          |          | 1         |           | 1         |
| <b>Total</b>   | <b>9</b>         | <b>12</b> | <b>7</b> | <b>19</b> | <b>1</b> | <b>5</b> | <b>3</b> | <b>1</b> | <b>21</b> | <b>13</b> | <b>91</b> |

Note: AA = African American site, AI = American Indian site, AN = Alaska Native site, AS = Asian site, GS = Generalized site, HI = Hispanic site, NH = Native Hawaiian and Other Pacific Islander site, and WH = Non-Hispanic White site.

Source: Comparative Ethnographic Studies CCM PI Analysis Sample Dataset

Coder/matchers coded up to three reasons for error for each of these interviews. Of the 91 coded sources, the five most frequently marked categories starting with the highest were mobility/tenuousness with 29, respondent confusion with 16, other, with 16, interviewer error with 10 and respondent concealment/refusal with 9. Language barriers were specifically tied to possible coverage error for three persons, and a hidden housing unit accounted for another three persons. The ethnographic reports can document the types of interviewer errors that occurred, how they were related to other sources of error, and identify how they may have contributed to possible coverage error. In order to understand how these chronically undercounted groups may be differentially impacted by cultural and other factors, the discussion now turns to cross-cutting themes across the ethnographic reports for the groups observed.

#### **5.2.4 Reasons for possible coverage errors: ethnographic report cross-cutting themes**

The ethnographic reports shed light on the possible sources for enumeration error separately from the records check study outlined above; we focus here on those sources of enumeration errors specifically associated with possible coverage error in Table 20. Comparing the results from the ethnographic observations, debriefings, and reports we identified several crosscutting themes paralleling the sources of error mentioned above: 1) interviewer error; 2) difficulty in gaining access to respondents; 3) language barriers; 4) mobility; 5) recall issues and 6) respondent reactions to multiple cold-call visits in 2010.

##### **5.2.4.1 Interviewer Error**

The first theme that goes to the heart of standardized interviewing methods is wide variation in interviewer behavior during their interviews, as documented by the ethnographers. To our knowledge, this is the first time that interviewer error has been identified as a source of possible coverage error. The ethnographers report a very wide range of interviewer behaviors, ranging from those who read questions exactly as worded all of the time as trained, to those who reworded or skipped most questions in the interview (Schwede, Terry, and Childs, 2012). Some interviewers routinely reworded questions to make them more conversational while a few interviewers deliberately shortened interviews to increase their productivity. Ethnographers observed that some interviewers shortened questions to the extent that they confused respondents, especially the foreign born. Some interviewers asked leading questions and seemed impatient when respondents did not reply immediately and sometimes continued without getting a respondent's answer. In addition to making major changes to wording or omitting questions, interviewers were also observed making errors by deciding on their own whether a person should or should not be listed, rather than adding the person to the roster, as instructed in training.

In most situations, interviewers seemed to react and modify their behavior to situations presented by respondents. In some situations, interviewers reworded or omitted questions to shorten the interview when they encountered respondents in the middle of doing something like eating dinner, perhaps to avoid break-offs. Interviewers also sometimes went off-script when attempting to interview angry or hostile respondents who may have been even more likely to break off or refuse. Ethnographers occasionally observed interviewers who were tired or rushing to finish at day's end who shortened, omitted or rearranged questions. Ethnographers also observed nonstandard behavior when a monolingual interviewer tried to conduct an interview

with a respondent with little or no English fluency or with a hearing-impairment. In these situations, interviewers may have felt the need to revise or reword questions to complete the interview. Leeman and Marsh (2013) elsewhere use the behavior coding method to document the extent of major rewording and omission of questions in the CCM, with 63 percent of their analysis tapes provided by the ethnographers on this evaluation. They documented that questions were read with no change or slight change half of the time and that 43 percent of the total sample involved major changes to question wording (Leeman and Marsh, 2012: 7).

In one Native Hawaiian site case, the respondent reported that the household was “in transition” around Census Day, but the interviewer recorded them as if they were living there. The ethnographer later learned during the debriefing that they had actually moved in after Census Day, and the owner could have been living there on Census Day. This CCM interviewer error appeared to have resulted in mis-enumerating two households.

Rewording questions, an interviewer error, was also noted by ethnographers to be a source of respondent confusion, another independently identified source of possible coverage error. In these situations, it was because the interviewer deviated from the question script or procedures that the respondent provided conflicting or inconsistent responses (coded as confusion), creating a possible coverage error that might have been otherwise avoided. In a case in the non-Hispanic White site, interviewer error connected with the instrument resulted in four possible outmovers being omitted from the roster.

#### 5.2.4.2 Difficulty Gaining Access

Ethnographers observed difficulty in interviewers’ gaining access to respondents. Difficulty in gaining access could contribute to missed addresses, hidden housing units, and could play a role in respondent refusal or concealment.

A hidden housing unit with three persons was in the African American site. It was on the same property, but separate from the main housing unit. Persons in that housing unit appeared to have been missed in the 2010 Census.

The African American site was in poor neighborhoods with a high incidence of crime. Many of the sample households were in multi-unit structures behind locked main entry doors through which one needed to be buzzed in by a resident. The ethnographer observed many soft refusals through the buzzer system, which made gaining access to these multi-units very difficult.

In the Native Hawaiian site, some respondents lived in housing units with fenced properties and locked gates in front of long driveways to maintain privacy. One interviewer went to an area labeled “dangerous” on the survey map and discovered a sign on one property giving a clear warning that trespassers were not allowed and threatening harm. The interviewer walked past the sign to the porch and fed treats to barking dogs. The owner appeared and yelled, “What are you looking for? That sign is out there. You should not come in here. Leave my property. Don’t come back!”

On the American Indian reservations, in rural Alaska and in the Native Hawaiian site, housing units were often isolated with no street names or numbers, making it difficult to find and verify



sample housing units with descriptions such as white house with grey trim or house near a specifically numbered telephone pole. Respondents in these areas had difficulty providing addresses or locations for other people. Even if one might know how to get to a different address, it would be very difficult to describe the house in enough detail for an interviewer to locate it later. This factor made describing living situations for mobile or tenuously attached household persons difficult, leading to possible coverage error by the inability to provide complete information on mobile household members.

#### 5.2.4.3 Language Barriers

The language barriers tied specifically to possible coverage errors in the records check identified the Native Hawaiian/Other Pacific Islander site where Tongan and Marshallese respondents had difficulty understanding the English language questions.

On both American Indian reservations, English is a second language for many respondents; most everyday discourse is in the indigenous languages. Many interviews were conducted in a combination of English and the reservations' indigenous languages because the observed interviewers were bilingual. During the interviews, it was often when topics of complex living situations arose that the discussions slipped into the respondents' native language. We note this here, not because it was tied to coverage error, but because it was observed as an adaptation to a potential barrier. There are two implications: 1) interviewers who speak the respondent's native language have an advantage; and 2) unless the instrument can be translated into other languages, data will be collected in on-the-fly translations, rather than through use of standardized translations, which is likely to introduce variation. Past ethnographies highlighted language as an enumeration barrier (de la Puente, 1993; 2004).

In both the Hispanic site and the Middle Eastern site, the ethnographers reported five contacts with their target population that did not lead to interviews during their visits because the monolingual English-speaking interviewers could not find a way to communicate with the respondent during the initial field contact. This was also the case in two Chinese households in the Asian site which could only be done with a bilingual interviewer. According to standardized procedures, the interviewers were to report these situations to a crew leader or a bilingual interviewer, either to have a bilingual interviewer accompany the regular interviewer or for the case to be reassigned to a bilingual interviewer. Some interviewers tried to muddle through the interviews in English, which was difficult for both parties, and could potentially have had implications for coverage. Others enlisted the help of other bilingual persons in the households to interpret. The authors think it is likely that these non-English speaking households would be revisited by bilingual interviewers at a later time in the CCM PI, but have no data on the extent to which this happened, because the ethnographers did not collect these names and did not provide the addresses and geocodes of cases that were not interviewed. This might be a topic for future research.

#### 5.2.4.4 Mobility

Mobility takes many forms that may influence census coverage. Many persons move permanently from one place to another. When the move occurred very close to Census Day,

these proved particularly difficult to count. Two persons in the non-Hispanic White site coded as possible census omissions moved into the observed housing units just before Census Day, and one possibly omitted person in the Native Hawaiian site moved out right around Census Day. Respondent confusion about whether persons moved in or out before or after Census Day was a source of error in several cases. Because possible coverage error could be attributed to multiple sources, some errors due to mobility were also reflected as interviewer error when the interviewer contributed to the problem. Several cases illustrate that recording the situation accurately may be especially difficult for both respondents and interviewers when persons move just before or just after Census Day.

Other cases of mobility involve “cycling,” which we define as going back and forth between two or more households on a regular or semi-regular basis. These situations have been characterized as “boomerangs” or “floaters” in past research and have been associated with hard to count groups (Bates and Gerber, 1998). In the American Indian sites, there were several cases of cycling. In one, a man went between households and it seemed he should be counted at the observed housing unit because he spent more time there than anywhere else. In another instance, a grandmother traveled back and forth to a distant city to spend time with her grandchild, though the exact amount of time spent and the place she should be counted were unclear. In the Generalized site, a young man went back and forth between a southern and a northern state; it was not clear where he should be counted. In the Alaska Native site, retired snowbirds moving among states were hard to count. (For information on snowbirds and other mobile populations, see Hunter (Childs), de la Puente, and Salo, 2003).

Again, ethnographers noted that respondents were sometimes confused as to whether they should mention persons with mobile or tenuous attachments to the household, and interviewers sometimes erred in deciding whether or not to include these persons on the roster. Past research has shown the connection with tenuousness and mobility and respondent uncertainty as to whether the person should be included as a household member (Bates and Gerber, 1998), but this study showed how interviewer error could exacerbate that problem by failing to roster someone that the respondent mentioned to a roster probe (as they were instructed to do in interviewer training).

#### 5.2.4.5 Recall issues and respondent confusion

Respondents’ ability to recall who was in the household specifically on April 1, 2010 decreases as the time elapsed from Census Day to CCM Interview Day lengthens. Ethnographers observed some of this in NRFU and it had become more of an issue in CCM PI observations more than four months after Census Day, especially for respondents who had in-movers and out-movers and who are asked a long series of questions to identify the alternative places and estimate the proportion of time spent in each of them back in March and April, and also over the last 12 months. For those cycling between two or more addresses, respondents had special difficulty in estimating the proportional amount of time spent in each. When respondents hesitated on this, there was a tendency for some interviewers to lead respondents by openly asking respondents if it was half time in each place. This is the most difficult scenario to resolve, because any equal time split precludes determining where one lived or stayed most of the time, so the default is where the person actually stayed on Census Day itself, April 1, 2010. Many respondents could

not answer this question definitively. April 1, 2010 is a typical day to those in the general population with no special salience except to Census Bureau employees. Had the interviewers not led with the suggestion of an equal split, respondents may have chosen different percentage distributions at each place that might have indicated where the person in question should have been counted. It would be well to move the CCM PI operation as close in time to April 1, 2010 as is feasible. Perhaps the move to Adaptive Design and faster automated processing of cases will make this possible.

#### 5.2.4.6 Respondent reactions to multiple cold-call visits in 2010

Over the course of the 2010 Census and CCM operations, some respondents expressed frustration with multiple cold-call visits from Census Bureau interviewers. Some CCM respondents in this late operation four to five months after April 1, 2010 reported two, five, and even ten cold-call visits from Census Bureau interviewers for 2010 Census and 2010 CCM interviews as well as Quality Assurance interviews. The Generalized site ethnographers and others said that some respondents were annoyed and angry because they had heard the earlier 2010 Census advertising that it was easy to respond—they just had to answer ten questions in ten minutes—and had complied but yet they were being repeatedly and unexpectedly recontacted by Census Bureau interviewers months after they thought the 2010 Census was over. The same was reported in the Native Hawaiian site and also heard occasionally in other sites. It was apparent from the second set of ethnographic reports on the CCM PI operation in the Generalized site and the Native Hawaiian site that some respondents in the CCM operation were less welcoming and less cooperative than those in the same general sites during the earlier NRFU operations when some enumerators were invited into houses to get out of the heat and sit down. This was made more difficult by the fact that the locations of CCM sites were not shared in order to preserve the independence of the CCM from the 2010 Census. As a result, local advertising could not be done to prepare potential respondents for additional Census Bureau interviewer visits. A small number of respondents who had been visited repeatedly openly questioned the capacity of the Census Bureau to handle data and keep track of enumerations. A few others mentioned government waste in multiple visits to collect the same data. For efficiency and cost purposes, as well as to try to maintain good relations with our respondents and prevent further erosions in response rates, it would be well to find a way to reduce the number of cold-call contacts we make with respondents in the 2020 Census.

### 5.2.5 Ethnographic site profiles: factors contributing to possible coverage error

In addition to the cross-cutting themes, there were also characteristics specific to each ethnographic site that could contribute to the identified possible coverage errors.

#### 5.2.4.1 Alaska Native Site (Donkersloot, 2011)

The Alaskan Native site includes Kodiak Island and nearby islands accessible only by plane or boat and part of the Alaska Peninsula. The one small city has large warehouses, a marina, a fish processing plant and a boatyard, as well as a Coast Guard base. Inhabitants are traditionally mobile, moving among home bases in relation to subsistence and fishing cycles. The ethnographer for this site identified mobility for work as a factor in possible coverage error.

Several examples demonstrate the complexity found in this site. One person was cycling between his home and his job on the distant oil fields during the fall and winter: six weeks in the workers' dormitory, then two weeks at home and back again. This man has a home where he and his family live, but he is away more of the time at the workers' dormitory. This case shows the difficulties of applying a residence rule to complex living situations. By counting him at home, the nuclear family and the man's contributions to his family can be identified. However, according to the rule of counting persons where they live and sleep most of the time, he should be counted at the dormitory, which would portray his household as a stem, not nuclear, family. In other cases, several men lived and slept on commercial fishing vessels for several months around Census Day. One of these men spends six to eight months a year aboard a vessel. It was not clear how to correctly count these men in the 2010 Census since this situation is not covered in the 2010 Census residence rule and situations document (Lamas, 2009). In the final example, interviewer error may have caused a respondent in the military to be miscounted because the interviewer inappropriately recorded a one-time overseas trip as a recurring monthly trip. In this site, primarily adult men were possibly affected by coverage error.

This profile shows how mobility for work in this particular geographic region could lead to omissions, as well as how interviewer error could exacerbate that problem. In addition, these situations were also shown to cause confusion for respondents, also leading to possible coverage error. The ethnographer also identified a hidden housing unit on the respondent's property behind the observed housing unit without a separate address with one resident. The housing unit and the person were missed by both the 2010 Census and the CCM.

#### 5.2.4.2 American Indian Sites (Fortier, 2012)

The two observed southwest American Indian reservations are among 562 recognized tribal entities and 314 reservations in the U.S. On both reservations, American Indian people learn their respective native languages before learning English. The native language is used in most conversations and in tribal governance and is an essential part of spiritual and religious identity.

Several factors may affect coverage in these sites. Many observed households were very mobile. Some persons moved back and forth between households on and off the reservation for work, others were tenuous and cycled among places, children moved among households, and a number of persons moved around Census Day. Determining where each person in these situations should be counted can be very difficult – some of these were coded as possible coverage error and some as undetermined. In one American Indian case, a baby of less than two months at the time of the 2010 Census who was identified in the CCM and considered a Census Day resident was not included in the final 2010 Census dataset included in the final localized 2010 CCM/Census matched dataset, though his mother was; because the possible omission occurred at a time other than when the ethnographer was there, no explanation is available for this discrepancy. Under these circumstances of lack of street names and numbers, remote and hard-to-reach residences, and frequent mobility, it is not surprising that respondents often had difficulty providing information for people who had moved out or were tenuously attached, making it harder to pinpoint where they should be counted and causing difficulty and confusion in the interviews.

A number of observed interviews went back and forth between English and the American Indian language. Because interviewers were recruited from the community, they sometimes were related to or acquainted with respondents and knew information relevant to the interview that was not always said aloud. Thus, this “insider knowledge” seemed very important in gaining respondent cooperation and clarifying situations where the respondents experienced mobility or had tenuously attached persons. The lack of city-style addresses in some areas and frequent mobility for work in distant locations or other reasons made it difficult for respondents to answer detailed questions on living situations for themselves or others causing them some difficulty and confusion in answering these questions. Because of these complexities and because of the limited English capability of some respondents, being able to speak the native language of respondents is a key to surveying these populations.

Another important cultural feature is the complexity of American Indian households and relationships, particularly those in tribal entities with patrilineal or matrilineal kinship systems. American Indian people in our reservation sites are related through matrilineal lines, reckoning descent from matrilineal clan affiliations. Generally, issues of family residence patterns and rules for clan exogamy and tribal endogamy constitute complex interrelationship patterns. These clans form the basic units of social cohesiveness, interchange, and are responsible for ceremonial obligations. These ritual obligations and the interrelationships between clan members can result in temporary household mobility. In a few cases, the ethnographer mentioned this mobility as a source of possible coverage error.

In this profile, because of the complexities on reservations with the geography, language and cultural differences, we observed the interrelation of mobility and respondent confusion and how they could lead to possible coverage error.

#### 5.2.4.3 African American Site (Lacy, 2011)

Chicago, Illinois has a large Black population with many characteristics previously associated with miscounts. The number of interviewers willing to go to the poor, economically distressed, high crime areas in this site after dark when residents tended to be home was small, particularly after a shooting occurred in the area a week earlier. Drugs were sold openly; in one case, an interviewer was advised by a respondent to stay away at a particular time due to a drug deal. Like in the 1990 ethnographic work in Harlem, this ethnographer witnessed “enumerator fear,” a factor previously associated with differential undercounting (Hamid, 1992).

This ethnographer identified several sources of possible coverage error: distrust leading to deliberate concealment of information; interviewer error; and mobility among the poor (Lacy, 2011). In two cases, female respondents appeared to conceal information about brothers living with them. In one, the woman identified her brother and said he had another address, but that she did not know it. She voiced concern that if she gave his other address, his status at college could be jeopardized. In the other case, the respondent’s brother was present prior to the interview, but left as it was starting. The ethnographer suspected that he lived there although he was not mentioned in the interview.

There were several situations involving children moving between households. A grandmother said her grandchildren did not live with her, but “stayed” there because she babysits them while their mother, who lives elsewhere, works at night. The interviewer had passed this house often, saw the grandchildren there most times and concluded the grandmother was concealing that they actually lived there. The 2010 Census residence rule is based on usual residence, where one *lives and sleeps* most of the time. In this case, the children may live in one place and sleep in another. Other cases demonstrated mobility of children staying with fathers or grandparents on weekends only. Respondents seem to have trouble deciding on the proportions of time spent in each place and ethnographers noted the interviewer error rewording questions to prompt respondents by asking if the child is “50/50” when they take a long time to ponder this. That is the most difficult situation to resolve because the question that follows requires knowing where the child was specifically on Census Day, which has been observed to be difficult for respondents to answer.

Two other cases evidenced mobility in different forms - a man currently living in an apartment was homeless on Census Day, and another man was in jail continuously from before Census Day to the interview day, but was nonetheless included on his fiancée’s household roster when he should have been counted at the jail.

Additionally, one situation involved three persons discovered in a hidden housing unit behind the main house on the property that appeared to have been missed by the 2010 Census but picked up by the CCM.

This ethnographer concluded that grandchildren and adult men were at the most risk of possible coverage error in her site. Distrust and concealment were strong themes of possible coverage errors for this group, as has been found consistently since the Valentine and Valentine (1971) ethnographic study. Though the Census Bureau has worked on outreach and communications campaigns directly targeting these issues, as evidenced here, the problem has not been completely solved.

#### 5.2.4.4 Native Hawaiian and Other Pacific Islander Site (Daniggelis, 2011)

The island of Hawai’i, or the “Big Island,” has many sparsely populated rural areas where Native Hawaiians comprise a larger-than-average population share. Native Hawaiians are disproportionately affected by poverty, represented as victims and offenders and, thus, are more likely to be mobile and hard to enumerate (McMillen-Wolfe, 2011). They are also much more likely to live in large households than the general population, also a prevalent characteristic of hard-to-count populations.

The Native Hawaiian concept of *‘ohana* is comprised of relatives by blood, marriage or adoption and is the fundamental unit in Hawaiian social organization. It comprises a network of kin that can and often does include persons living in other housing units. This is the core economic unit that provides social and economic support for household members and other relatives. One instance of possible coverage error due to mobility was within an *‘ohana*. This was a situation where three persons had moved just after Census Day, but had been living in an “*‘ohana* housing unit” on Census Day. According to records, they seem to have been miscounted at the sampled housing unit.

In this site, not adhering to culturally appropriate behavior was also identified by the ethnographer as a barrier to enumeration. She identified two Hawaiian cultural values—*pono* (respectful in every way you can be) and *ha`aha`a* (humility)--as being important to follow while interviewing Native Hawaiians. Some interviewers did not routinely practice this. When interviewers rushed through interviews, shortened questions, or were impatient, problems occurred. The link between not following these cultural practices and “interviewer error” documented in Table 20 is shown in the following case. A possible omission occurred when the interviewer did not listen attentively and failed to roster a daughter-in-law that the respondent mentioned had lived there. During the debriefing, the ethnographer verified this and learned the daughter-in-law had since moved out, but should have been counted there on Census Day. In another case, the ethnographer observed a White interviewer who was domineering, cut respondents off, and made jokes about Hawaiian names, thus “discouraging” a respondent from reporting that he was part Hawaiian. This type of behavior can lead to respondent concealment or refusal. Politeness and respect have also been noted to be valued by Koreans (Kang, 2006), Chinese (Pan and Lubkemann, 2012), and American Indians (Fortier, 2012).

Other instances of mobility involved determining where persons with multiple residences should be counted. Examples of persons splitting time between two residences included a grandchild going back and forth between grandparents’ homes; a mother giving contradictory information about whether her children should be counted; a college student living away; and adults living away to get their high school diplomas but going back and forth to visit their parents on weekends.

In this site, nearly all sources of possible coverage error were observed to some degree. Mobility was the predominant factor, but interviewer error, language barriers and respondent concealment and refusals were also seen. We suspect that the lack of cultural etiquette by the observed interviewers only compounded the difficulty enumerating this mobile group. Additionally, a married couple in the Native Hawaiian site who in the CCM interview gave every appearance of having lived in the same housing unit with no other place was not found in the final local 2010 Census records included in the 2010 CCM/Census matched dataset. In another case, no information was available as to why just five persons of eight listed in the CCM as having lived there on Census Day were found in the final local 2010 Census dataset matched to the CCM. Those that seemed to be missed at this housing unit in the 2010 Census included the uncle and two of the respondent’s young biological children, one of them less than a month old.

In addition, three persons refused to be interviewed, with two of them saying they were opposed to recognizing the federal government and participating in the 2010 Census. It is not clear if these refusals were in response to public calls in the media by Hawaiian Sovereignty Movement activists to boycott the 2010 Census.

#### 5.2.4.5 Hispanic Site (Newby, forthcoming)

The Dallas/Fort Worth/Arlington Metroplex area is the second-fastest growing region in the US and the city of Dallas has a large and rapidly growing Hispanic population comprised of US-born Hispanics and immigrants (Garcia, 2012). Of the 2010 population of 6.4 million, Hispanics of

any race constituted 28 percent in 2009 (U.S. Census Bureau, 2010). It has a very large, diversified economy, with farming, ranching, oil, and energy industries. Individual poverty rates in Metroplex cities are above the U.S. rate of 14 percent, and vary widely. Arlington has a poverty rate of 14 percent, with Fort Worth at 17 percent and Dallas at 22 percent (American Community Survey 2005-9 estimates).

Distrust of the government and suspicion of the CCM interviewers were noted by the ethnographer. Some respondents were suspicious because they had heard broadcasts from Spanish media Univision that the 2010 Census was already completed. This was also a time of heightened tensions about immigration policies in Texas and in Arizona and some unknown number of respondents may have been illegal immigrants who were anxious about participating and providing information to the government that they may have feared might be shared with immigration officials, thereby risking deportation.

Language barriers have already been mentioned as a source of possible coverage error in Table 20 for completed interviews. But language barriers also prevented some interviews with non-English-speaking respondents from being conducted during the ethnographer field observation study periods. This was demonstrated in the Hispanic site where potential interviews at five housing units could not be conducted because the interviewers could not speak Spanish and did not find a bilingual person nearby to serve as an on-the-spot translator. The ethnographer's role was to document what happened during the interview situation and conduct debriefings – it was not her role to intervene as translator. In these cases, the interviewers said they would either call their crew leaders or ask for the case to be transferred to a bilingual interviewer, or call a bilingual interviewer directly to schedule a later interview. In one of those cases, she was able to observe the bilingual interviewer conduct the interview. In one of the other cases, the interviewer was able to make an appointment for a return visit with the respondent, but that respondent was not home at the agreed-upon date and time. In another two interviews, the English-only interviewer managed to conduct the interview in English with a primarily Spanish-speaking respondent, with difficulty for both parties. She was able to accompany two bilingual non-native Spanish interviewers and observe interviews in Spanish. One interviewer was Argentinian and his Spanish was somewhat different from the Spanish spoken by the primarily Central American Hispanic immigrants in that location. She recommended increasing the number of bilingual interviewers and matching the origin of bilingual interviewers to the origin of local respondents. The shortage of bilingual interviewers in this diverse area increases the time and costs of interviewing non-English-speaking respondents.

In case Hispanic site case, a woman was living with her unmarried partner and gave no indication of having stayed anywhere else over the past year, but was found during CCM matching to have been duplicated in another housing unit in the 2010 Census, according to the 2010 CCM/Census matched dataset. The CCM outcome was to count her at this housing unit. This case has a missing value for whether she was counted here in the 2010 Census.

#### 5.2.4.6 Middle Eastern Site (Foster, 2012)

Interviews were observed in five communities in Detroit, Dearborn and elsewhere in Wayne and Macomb Counties. The mortgage crisis, automotive industry crisis, and the economic recession



had a profound impact on Michigan and on Detroit in particular. Wayne County unemployment consistently exceeded the state average, at 16 percent in August 2010 down from 18 percent a year earlier. The Detroit/Dearborn area was chosen due to having the highest proportion of this minority in the US. Persons of Arab ancestry comprised nine percent of Dearborn, a 21 percent reduction from the Census 2000 estimate.<sup>19</sup> Two communities had large Middle Eastern presence, with mosques and Islamic centers, Arabic-language signs in calligraphy, halal grocery stores, advertisements for Al-Jazeera and Al-Arabiya and many Middle Eastern restaurants.

However, none of the five accompanied interviewers was Arab or other Middle Eastern or spoke Arabic or Chaldean. As in the Hispanic site, the language barrier prevented English-only interviewers from conducting interviews in five housing units with markers at their residences of Middle Easterners. In five cases, interviewers either asked her to translate questions to identify bilingual household members who could translate or to help respondents having difficulty in understanding and answering in English. She declined politely; her role was to document how such situations were handled. She observed interviews in households with Middle Easterners.

Faced with an elderly Middle Eastern respondent who could speak no English, an enterprising monolingual English-speaking interviewer obtained the name and phone number of her son from a neighbor and arranged an interview with the son at his workplace. The respondent's cousin served as the on-the-fly translator, which likely led to some major wording and possible meaning changes. While the interviewer's initiative in this case did get the interview completed, it would have saved time and costs and likely increased data quality for a trained bilingual interviewer to have made the first contact and conducted this interview.

Wider cultural/religious and political factors decreased the ability of the ethnographer to observe at least half of her interviews with her target population of Middle Easterners in this site in late August, 2010. The first was the coincidence that the CCM data collection period from mid-August to October overlapped to a large extent in 2010 with Ramadan, the one-month religious period during which Muslims meditate and fast from food and drink from sunrise to sunset. The timing of the interviewer and ethnographer visits in the late afternoon and early evening conflicted with religious obligations and practices of Arabs. Some Arabs did not answer the interviewer's knock on the door because they were preparing or eating the evening meal breaking the fast of 14 or so hours. Some interviewers were not aware that Ramadan had started; cultural awareness training would have helped. Partly due to this, most of her Middle Eastern interviews were with Chaldeans and Lebanese, just a segment of the Middle Eastern population in this area. This cultural/religious factor could also have potentially affected differential coverage of Arab and other Muslims if a larger proportion of Muslims were interviewed after multiple visits or late in the data collection phase or missed. It was unusual that Ramadan fell in August, 2010 during CCM PI; Ramadan is based on the lunar calendar and moves up about two weeks each year.

Current events during 2010 also had possible effects on coverage error in terms of Middle Easterner willingness to be interviewed and refusal rates. According to the ethnographer, there was tension in the community surrounding the Middle Eastern site as September 11 approached, due to the plans to build a mosque close to the former Twin Towers site. One respondent

---

<sup>19</sup>[http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS\\_10\\_5YR\\_DP02](http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_10_5YR_DP02)

insisted on photographing the interviewer and the ethnographer and consulted by phone with relatives as to whether to share any information with them. Another refused. Another potential factor for concern was the national call to boycott the 2010 Census by Middle Eastern activists pushing to get a Middle Eastern category added to the race question, though there is no way to measure possible effects of that in this evaluation.

A factor that was not specifically linked to possible coverage error in this study, but could affect coverage measurement through weakened ability to identify, detect and resolve duplications is differences in Arabic and English writing and spelling systems that can produce inconsistent transliteration of Arabic names when translated into English (e.g., Muhammad, Mohamed, and Mohammad). This is due to the Arabic alphabet having letters/sounds not used in English and Arabic writing that typically does not show soft vowels. While the particulars are different, this transliteration issue potentially decreasing the ability to unduplicate cases and thus raising the risk of undetected duplications is the same for Chinese, as discussed in Section 5.1.5.6. This could be a factor in differential coverage of these groups. This is a caution, too, for those who are assessing the feasibility and accuracy of administrative records for census enumeration or coverage purposes.

Just one observed Middle Eastern household had a possible omission coverage error. The young female respondent stated clearly that she had lived in the household with her family on April 1, 2010 and continuously through the August interview date and did not stay anywhere else. Her name was the only one from the family that was not included in the final 2010 Census dataset included in the final localized 2010 CCM/Census matched dataset. There is no census source that explains her absence from the local final 2010 Census dataset.

#### 5.2.4.7 Asian Site (Schwede, forthcoming)

San Francisco is racially and ethnically diverse with relatively high proportions of Asians (33 percent) and persons of two or more races when compared to the national average (5 percent and 3 percent, respectively). The Bay area is a cosmopolitan urban area. Tourism, finance, and technology are the top drivers of the economy; the tourism and other sectors were hurt by the economic recession. Several of the interviewers were highly educated professionals who were available due to the difficulty in finding permanent jobs.

In contrast to the NRFU/UE observations which were in the homogeneous Chinese enclave in Chinatown, the CCM sample cases were dispersed among the general population in the San Francisco Bay area, with none in Chinatown. Chinese and other Asians living dispersed within the wider racially and ethnically mixed population were more likely to speak English. The challenge was to identify pockets with larger proportions of Chinese. Fortunately, the CCM manager was knowledgeable about this and paired the ethnographer with interviewers going to those areas, although any potential respondent's race would be unknown until the resident answered the race question. Seventeen of the 41 interviews were with Asians: 11 with Chinese and the remaining six with Indonesian, Filipino, Japanese, Vietnamese and Korean, and another Asian.

None of the interviewers accompanied appeared to be Asian or spoke an Asian language and the language barrier was noted to some extent. The 17 Asians respondents ranged from fully fluent and born in the U.S. to those who appeared to speak and understand virtually no English. Fifteen of the interviews were in English. In eight of these, the interviews went well; in four, another person in the household (usually a child) translated or answered; and in one, a visiting son was his mother's proxy. Three of these five were Chinese, while one each were Japanese and Vietnamese. The remaining two English-only interviews were a strain on interviewers and respondents to understand, and would better have been done in Chinese. However, enough information was obtained to decide whom to count in these homes.

The difficulties of trying to complete CCM interviews at two sample housing units where respondents spoke no English illustrates problems language barriers cause. In the initial visit, the monolingual English-speaking interviewer could not even communicate enough with the elderly woman to get the respondent to look at the language identification card. With no one else around as a possible translator, the interviewer and ethnographer had to leave without establishing what language she spoke, though she appeared Chinese. No bilingual Chinese interviewers had been hired in this area yet by the regional office in Washington, so one had to be requested from 40 miles away to serve as the interpreter for the monolingual interviewer. After a brief conversation with the same elderly woman, the interpreter said she could not communicate well because the respondent spoke Cantonese which she, a Mandarin speaker, could not fully understand. It was only through the help of a trilingual mailman that this interview was completed. He both allayed the respondent's fears of providing data to the government and translated hard questions between Mandarin and Cantonese.

After the mailman left, the team tried to interview another Chinese woman, but she was also a Cantonese speaker. The bilingual interviewer resorted to translating each question and writing the Chinese characters on a piece of paper so that the respondent could read them and give short answers. At the end of the interview, the bilingual interviewer asked this woman to accompany us and establish contact with Chinese neighbors. This woman agreed and the elderly woman convinced another neighbor to do the interview. In that case, the bilingual Cantonese/English teenage daughter served as the interpreter. Later, the teenager agreed to help in other interviews on her block.

These three cases with monolingual Cantonese-speaking respondents demonstrate several important points. First, in some cases, monolingual English-speaking interviewers cannot communicate at all with non-English-speaking respondents and some recontact by a bilingual interviewer is necessary, taking more time and raising costs. Second, sending any bilingual interviewer is not enough in cases of the Chinese and other groups who have more than one language. Third, it is important to either know what language the respondent speaks or send an interviewer who can speak both Cantonese and Mandarin. Fourth, some non-English-speaking Chinese were distrustful and concerned about providing data and needed information to allay their concerns. Fifth, having a neighbor or a community leader serve as cultural liaison while escorting the interviewer made a large difference in convincing others to participate in the CCM interview. Sixth, it was learned that the Language Identification Card used to help respondents identify their languages was not of much use for Chinese speakers. The two entries on the card just asked about preference for traditional or modern script; they did nothing to identify the

spoken language of the respondent, which is critical to learn in order to assign the appropriate bilingual interviewer. The Chinese entries on the language card require revision for field use.

The one case of possible coverage error in this site was a possible erroneous enumeration that shows a mismatch between “usual residence” defined by the residence rule as where one lives and sleeps most of the time in reference to Census Day, and where one belongs long-term, as subjectively defined by the respondent. The resistant respondent appeared to be non-Hispanic White (he declined to give his race). His son had been away at college and graduated and then had gone overseas for an internship, returning to his father’s house in mid-April. Technically, the son had spent most of the seven months preceding Census Day away from home, so his father’s house was not his usual residence. Being out of the country for an extended period around Census Day renders him out of scope for the 2010 Census, but he was counted here according to the Census results in the final localized 2010 CCM/Census matched dataset.

#### 5.2.4.8 Non-Hispanic White Site (Hubbard, 2011)

Interviews were observed in three counties east of Kansas City, Missouri with a combined population of 995,575 persons (U. S. Census Bureau 2010 data). All three counties are predominantly non-Hispanic White: Clay County (84 percent), Cass County (90 percent), and Jackson County (63 percent) (Census 2011 Population Estimates Program). These counties are a mix of rural, small town, and urban settlements, with cities ranging from 10,000 to 100,000 people. Farms produce corn, wheat, and feeder cattle, supported by distributors and retailers of those products.

In one interview in this site referred to earlier, interviewer distraction led to four outmovers being possibly omitted. The interviewer asked an inmover the key outmover screener question on whether anyone else had lived or stayed here in March or April and was no longer here. The respondent mentioned the roommate’s wife and children but the interviewer was distracted and did not pick up on this. He said, “I must have pushed a button too fast or...alright, we’ll just go on.” It seems the interviewer inadvertently hit “no” to the screener question, because the instrument skipped directly to the screen asking for nicknames, not activating the important instrument path to collect names and data on outmovers. As a result of this interviewer error, four outmovers who perhaps should have been counted in the housing unit were omitted from the CCM. The ethnographer collected those names and some demographic characteristics in the debriefing. The respondent’s roommate’s wife and three children had lived here but were in the process of divorcing around April 1, 2010. One day the wife left suddenly with the children, without prior notice and without leaving a forwarding address. These four were not found in the final 2010 Census dataset included in the final localized 2010 CCM/Census matched dataset. The team could not say for sure if they were or were not omitted from the overall 2010 Census since the search area in the custom PerMaRCS dataset was limited geographically. The possible omissions here in the CCM were due to interviewer error and possible issues with the laptop and/or the automated instrument. In the 2010 Census, the possible omissions were likely due to the living situation of a pending divorce and separation.

The lack of respondent knowledge was a factor in the possible omission of one Census Day renter. In this case, the inmover could provide only the first name of the former resident of her

apartment who had moved out shortly after April 1, 2010. She became angry when the interviewer continued to ask questions she could not answer about that outmover. The outmover was not on the final 2010 Census roster and the CCM was not able to match the person elsewhere with just a first name, using the final localized 2010 CCM/Census matched dataset. There were cases in other sites like this, where new in-movers were unable to provide information on prior out-movers who should have been counted there or to say with certainty whether the housing unit had been vacant or not on Census Day.

Hence, interviewer distraction, technical problems, sudden mobility due to family breakup and divorce around Census Day, and an unknowledgeable respondent were some of the reasons for possible coverage error in this site.

#### 5.2.4.9 Generalized Site (Sando, 2011)

Broward County was selected as a Generalized site because, in the Census 2000 sample tract-level data, it had roughly the same race/ethnicity distribution in Census 2000 as the nation had, according to population estimates in early 2010 (i.e., 70-75 percent non-Hispanic White, 10-15 percent African American, and 12-18 percent Hispanic). Further, we selected this site for its average Census 2000 “hard-to-count” score in the Census Bureau’s Planning Database (special computer run by Nancy Bates). According to the 2010 Census, the county had become more racially and ethnically mixed over the decade: non-Hispanic Whites (44 percent); African Americans (26 percent); and Hispanics (25 percent) with the rest distributed among other groups. Florida has a large tourist economy, retirement population, and many “snowbirds” from Canada or New England who arrive in the fall and stay until spring. As a vacation and retirement destination, the state experiences periodic boom-bust economic cycles. The ethnographer explains that in 2007, the nation began moving into severe recession, stemming from multiple factors and this county was hard hit; in January 2010, this county had the highest foreclosure rate in the state: one in every 105 homes (*Sun Sentinel, 2010*).

In this site, a number of situations had to do with job loss and change of residence, a hidden housing unit and an in-mover not able to provide information on an out-mover. In the hidden housing unit case, the respondent claimed to have included a man on her mailout census form who was living in an apartment off the garage on the same property to ensure he was counted in the 2010 Census, even though he was away a lot. If he were there most of the time, he should have been counted in his own garage housing unit. The final 2010 Census dataset included in the final localized 2010 CCM/Census matched dataset showed him counted incorrectly in her housing unit as part of her household. He was actually the owner of the house and garage who had lost his job and was renting the house to the respondent. He was thus included in the 2010 Census, but in the wrong housing unit, giving the false appearance that the household included a nonrelative.

A coding or scanning problem during the processing of 2010 Census forms may have been the reason why one young cycler may have been omitted from the OHU in the final 2010 Census dataset included in the final localized 2010 CCM/Census matched dataset, though his name had appeared on the scanned 2010 Census form associated with this household. This was the young

man who was cycling between this place and one in another state. As a result, this man was picked up in CCM but not matched to the 2010 Census record for this or any other housing unit.

One additional theme identified by the ethnographer from this and the Native Hawaiian site was respondent frustration, anger, and resistance to multiple visits by Census Bureau enumerators or interviewers. A number of respondents stated they had already been contacted as many as four, five, and even ten times and many were fed up with this. Many insisted they had already completed and mailed in their “ten questions in ten minutes” as widely advertised and they could see no reason why they were being contacted again. Some felt harassed and some took this as a sign of inefficient government practices that wasted tax payer money. Others were afraid the interviewers were scammers and did not find the temporary paper badges to be convincing. The authors suggest giving interviewer badges with their pictures on them for the 2020 Census. They also are fully in support of 2020 Census research on reducing and improving followup visits to the extent they do not compromise data quality.

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

### **6.1 Census-related Evaluations, Experiments and Assessments**

This evaluation has close links to the NRFU Behavior Coding Evaluation (Childs and Jurgenson 2011). As noted, 193 (95 percent) of the 204 tape-recorded NRFU interviews behavior coded in that evaluation were collected by this evaluation’s ethnographers. The Section 5.1 write-up of observed interviewer and respondent behavior complements the results of behavior coding of interviewer and respondent responses across many interviews. They found that enumerators read questions as worded in just 37 percent of all questions. This baselines the frequency of question rewording and omissions, documenting the need to improve training to emphasize the importance of reading questions as worded.

Because that evaluation used the tapes from our household interviews, there is a natural household link between the cases in these two evaluations that allowed us to merge the two datasets into one and explore demographic characteristics of households, persons and even some characteristics of interviewers that are linked to behavior coding. We co-authored a presentation on the interaction of interviewer and respondent race and behavior, as documented with behavior codes (Terry, Childs, Schwede, and Ryan, 2012). This joint dataset will provide many opportunities for research in the coming years with the linked ethnographic/behavior-coded cases.

Also, this evaluation included some households in which level of respondent fluency in English and other languages was an issue. Those cases complement the findings in the “Observing Census Enumeration of Non-English-Speaking Households” Evaluation (Pan and Lubkemann, 2012). In that evaluation, foreign-language-speaking ethnographers were paired specifically with bilingual enumerators to enumerate Chinese, Middle Eastern, Russian, Arabic, and other foreign language-speaking respondents so that all parties—the enumerator, the respondent, and the ethnographer—could understand and speak with each other. In contrast, our ethnographers accompanied any interviewers and reported on the problems and outcomes experienced when one or more of those parties could not converse in and understand the respondent’s language.

That presents a very different dynamic in the interview than when all parties speak the same language (Schwede and Terry, 2012a, Schwede, 2011a).

Further, that non-English-speaking evaluation included observations of NRFU operation interviews with Middle Easterners in Michigan. In this evaluation, the Middle Eastern site during NRFU had to be dropped for lack of an available Arabic-speaking ethnographer. Transcripts of their 300 or so NRFU interviews, both for the Middle Easterners and the other race/ethnic groups, could be compared to the roughly 300 transcripts from the authors' tapes collected at the same time in our other sites, producing a dataset of more than 600 cases oversampled for race/ethnic and non-English-speaking respondents in the 2010 Census. That could be a dataset available for ongoing survey methods research on interviewing respondents in hard-to-count populations.

The NRFU component of our evaluation also complements the "Avoid Followup Experiment" in the 2010 Census Alternative Questionnaire Experiment mailout tests on persons at risk of census duplication. That experiment tested the X13 2010 Census overcount booklet with a new experimental sequence of followup questions for persons who are recorded by respondents on the form as "yes" to "sometimes live or stay somewhere else" that had shown promise in cognitive testing of the experimental question sequence (Schwede, Sorokin and Yelei, 2009). For those persons with a "yes" answer to the possible overcount question, the experimental X13 form asked for the alternative address. It also asked where the person lives and stays most of the time and where the person was staying on April 1, 2010. The results indicated that "people were willing to provide an address when prompted in their initial (mailout) census questionnaire and their complex living situation and duplications could be resolved using the current residence rule without an expensive followup interview" and could reduce recall bias. That booklet had a higher response rate than one-page census forms (Jackson and Wechter, 2012); that approach looks promising.

Debriefings from this comparative ethnographic evaluation provide qualitative data on living situations of persons who may be counted in more than one place that may help to explain how and why census some duplications occur. These findings can complement those in the "2010 Census Effectiveness of Unduplication Evaluation Report" (Heimel and King, 2012).

Debriefings in this comparative ethnographic evaluation provide more information on how and why persons may be duplicated and who is affected. This can provide useful background information on factors within the interviews that may affect duplications in general.

Other associated assessments include the "2010 Census Update Enumerate Operation Assessment Report" (Fallica, Heimel, Jackson, and Zhang, 2012) and the "2010 Census Nonresponse Followup Operations Assessment" (Walker, Winder, Jackson, and Heimel, 2012).

## **6.2 CCM-related Evaluations, Experiments and Assessments**

The CCM component is also related to and can provide supporting data to another Census Program for Evaluations and Experiments (CPEX) evaluations and to one CPEX experiment, as mentioned earlier. The CCM component of this evaluation has close links to the CCM Behavior

Coding Evaluation (Leeman and Marsh, 2013). About 172 (63 percent) of the 271 tape-recorded CCM interviews behavior coded in that evaluation were collected by this evaluation's ethnographers. Our write-up of observed interviewer and respondent behavior in Section 5.2 complements the results of behavior coding of interviewer and respondent responses to individual questions across many respondents and many interviewers. Their findings that questions were read with no change or slight change half of the time and that 43 percent of the total sample involved major changes to question wording (Leeman and Marsh, 2012: 7) provide numerical documentation of interviewer and respondent behaviors during the survey process. These can be used to develop guidelines for training for the 2020 Census and substantiate the need to have some system of better ongoing monitoring of interviewer behavior in the field.

Consideration could be given to matching the Leeman/Marsh behavior coding results to the corresponding household data in this evaluation's analysis sample. This would make it possible to do much more in-depth analysis of the relationships among coverage, and interviewer good and not-so-good behaviors. Linked files would also make it possible to examine the distributions of interviewer error by interviewer across household interviews, to see if interviewer error is confined to a small number of interviewers or if it is more widespread.

## **7. Key Lessons Learned, Conclusions, and Recommendations**

### **7.1 Census Component**

#### **7.1.1 Records check: numbers, types and reasons for inconsistencies**

In the NRFU/UE component of this evaluation, 103 (13 percent) of 786 persons were found to have inconsistencies among the four data sources in the records check stage: the standard EQ interview, the ethnographer assessment, the final localized 2010 Census unedited dataset, and the coder team assessments. These 103 persons were in 44 (17 percent) of 257 housing units.

The sources of inconsistencies documented most frequently in the statistical analysis include: interviewer error; mobility/tenuousness; respondent concealment/refusal; address missed in the 2010 Census; respondent confusion; language barriers and other factors. African Americans and American Indians/Alaska Natives (almost all on the American Indian reservation) across sites had higher proportions of inconsistencies than other groups, though there is nothing inherent in being in these groups that causes such inconsistencies. The African American site and the American Indian site had higher proportions than the other research sites. There appeared to be no consistent gender differences but the youngest age cohort ages zero to four had the highest proportion of inconsistencies per age group. While the results of this small-scale qualitative study are not directly comparable with nationwide statistics, this same age cohort also had the highest undercount in a national study (O'Hare, 2013). Some case situations of young children being left off rosters even when present and visible during the interviews were described in the Hispanic and Native Hawaiian profiles; other cases of children cycling between and among households were presented in the African American, non-Hispanic White, and other profiles. Some of these were formal joint custody but others were informal arrangements of child sharing or visiting.



### **7.1.2 Reasons for record inconsistencies in ethnographer report cross-cutting themes and profiles**

From the qualitative data across sites, five cross-cutting general themes associated with inconsistency were identified: enumerator error; difficulty gaining access to respondents in urban and rural areas; language issues both within completed interviews as well as in interviews that could not be started without the assistance of an interpreter either within the household or in a later visit by a bilingual enumerator; cultural differences and issues; and many forms of mobility.

The qualitative data from actual inconsistent cases within the context of each ethnographic site illustrate the sources of error raised in the statistical analysis and also raise issues for consideration. In the Alaska Native site, mobility for long-distance work on land and sea in the Coast Guard, Merchant Marines and on the North Slope oil field raised issues of potential overlap in housing unit and group quarters operations. Other factors in inconsistencies and possible miscounting included multiple types of interviewer error, local and inter-state cycling among households, and a man's attempts to avoid the 2010 Census.

In the American Indian site, vague descriptive addresses given due to the lack of street names and addresses contributing to inconsistencies and possible coverage errors were linked to 2010 Census maps for remote areas that had few or no land forms or topographic details and were not uniform in scale. Inconsistency was also linked to mobility and a local enumerator who knew the family appeared to miss the significance of the respondent's statement that one child was living elsewhere with her father to attend school.

In the Native Hawaiian site, the same issues of lack of street names and addresses were linked with 2010 Census maps that were difficult to follow. In addition, at least one "hidden" household on a large rural property was discovered and enumerated in the field but appeared not to have been included in the localized dataset. Question rewording that dropped reference to April 1, 2010 was also an important factor, as was a respondent's interpretation of the question as being about who was included in his family, rather than who lived or stayed there on Census Day.

The Asian site, too, demonstrated the issue of vague addresses and inconsistencies but in urban multi-unit buildings. Here, irregular housing arrangements were found with two unrelated families sharing one housing unit but not functioning as a social household and conversely, one family using three contiguous SRO-type rooms as one housing unit. These cases raise the factor of mismatches of respondent conceptions of who "belongs" in their "households" and residences with the 2010 Census definitions of the "household" as all who share one physical housing unit, whether related or not. Other factors in inconsistencies included language and spelling barriers, and bilocality. This site also demonstrated the range of Chinese languages and dialects and the critical need for bilingual enumerators from the community to do on-the-fly translations when standardized translated EQ forms are not available.

The same need for bilingual enumerators to do on-the-fly translations due to the lack of a Spanish version of the form was an important factor in the Hispanic site as well. In that site, question rewording, household complexity, and respondents who were unknowledgeable or who provided inaccurate or incomplete data to comply with the interview but also offered resistance or for other reasons were other factors associated with inconsistencies. A high soft refusal rate in this site also suggested concerns about confidentiality and fears of deportation for some during the highly charged debate about strong anti-immigration laws at that time.

As in the Hispanic site, respondent concealment and distrust of the government were found in some inconsistent cases in the African American site. Other major factors in inconsistencies were frequent major question rewording by enumerators, mobility, and some housing units and persons that appeared to be missed in “Smith/Jones” type situations. Two of these cases appeared to be separate basement housing units that were enumerated in the field but appeared not to have been in the localized final 2010 Census unedited dataset, raising the suggestion of possible issues in processing more than one form for the same physical address that may have more than one housing unit. In some of the cases of inconsistencies found in this study that were linked to major enumerator error, the persons were still counted in the final 2010 Census dataset, despite the lack of adequate substantiation in the interviews themselves. As stated earlier, inconsistencies in this analysis do not necessarily imply possible coverage error in the 2010 Census.

In the non-Hispanic White site, a “Smith/Jones” situation that raises questions about the quality and accuracy of proxy reporting in complying with the interview but possibly providing inaccurate information was documented in a low-income housing project in the non-Hispanic White site during a blitz operation. Additionally, the situation of mobility of children cycling between housing units and trying to determine from the time spent in each where they should be counted was a reason for inconsistencies. Cycling cases can be difficult to resolve.

In the Generalized site, the important issue identified was enumerator error in not always following up with in-movers to make sure they were included in the 2010 Census at their prior April 1, 2010 places.

### **7.1.3 Research implications and recommendations**

#### **7.1.3.1 Enumerator error**

Enumerator error in the form of rewording and omitting questions, leading respondents, not listening fully to their answers and not following standardized procedures for assuring that in-movers and persons in WHUHE arrangements are counted at their other residences cuts across these eight sites and varies among interviewers. This can be partially addressed through survey research on the questions and procedures as well as through research on enumerator behavior and training.

On the survey research side, the authors suggest reviewing and revising key coverage questions S2, S3, S5, and roster question 1 to emphasize the April 1, 2010 reference date and to break up double-barrelled questions in S2 and S3 and streamline question S5. Some very informal cognitive testing of options such as “As of April 1, 2010,” “Around April 1” and “On April 1”

might be reviewed as very preliminary background findings (Schwede, Sorokin and Yelei, 2009). The authors also suggest breaking the long lists of response categories in overcount flag question 7 and omission flag question H1 into shorter questions more easily read by enumerators and likely more readily understandable to respondents. The authors recommend exploring the possibility of adding the question sequence from the Avoid Followup Experiment X13 form to the Enumerator Questionnaire for NRFU and for UE (Jackson and Wechter, 2012).

The authors suggest that enumerator error also be addressed through improvements in enumerator training. They suggest developing and testing a short training module that describes the processes the Census Bureau uses to develop and test questions and why it is important for enumerators to read questions as worded, so that enumerators understand their vital role in the process of collecting valid and reliable data. The module could include some discussion of what each question is intended to measure and how it is reflected in final statistics. The module could use examples culled from this or other research to show how even seemingly small changes like shortening “living or staying” to just “living” or leaving off “April 1” at the end can result in some persons being left off a roster who should be counted there, or including some who should not be.

The authors suggest developing and incorporating role playing to try to establish rapport and engage difficult respondents or those with limited English or disabilities, if this is not already being done. This could help enumerators become familiar with these types of situations and with how the questionnaire wording might be adapted without changing the meaning.

Additionally, it would be useful to supplement enhanced training with some method of more frequent monitoring and feedback on enumerator behavior, through both more frequent crew leader observation and also through Computer-Audio-Recorded Interviews (CARI) or other interview recording mechanisms. Interviewers could then be held more accountable on a regular basis for the way they present questions to respondents, not just when an observer is present. Knowing that at least some of their interviews were being recorded and might be reviewed might reduce some of the rewording and omission of questions.

Finally, the authors document that a handful of some of the “best,” “most productive” enumerators may be deliberately shortening and combining questions as a means to complete more interviews per area or day. This suggests a possible issue with the enumerator incentive structure that may be a little more oriented toward quantity, than quality, of interviews and perhaps might benefit from review.

### **7.1.3.2 Cultural competency and cultural facilitators**

Another issue that was raised in this report was the ability to interview persons of diverse backgrounds and languages. Some of the ethnographers gave examples of how enumerators were unfamiliar with basic rules of pronunciation of names in Spanish or of appropriate ways to act with Native Hawaiians, or of differences among certain populations within their geographical areas. A few of the ethnographers mentioned an occasional enumerator who was culturally insensitive in some way, more often in the context of rewording or leading respondents.

The authors note that local field managers in a number of the research sites did succeed in their aim of hiring enumerators and crew leaders from the local community to conduct interviews there, particularly in the African American site, the Chinese site, and on the two American Indian reservations. For the most part, this was highly beneficial, but occasionally prior familiarity with a respondent and his/her living situation led the enumerator to probe more intrusively, as in one case summarized in the Alaska Native profile, or to overlook a respondent's statement that one of the children she rostered was actually living elsewhere for the school year.

Another recommendation is that in sites where it is difficult to hire local people of some minority groups, such as the Native Hawaiians and Marshallese on the Big Island, the practice employed in Hawaii of hiring well-known and respected members of those small race/ethnic communities to serve as cultural liaisons be adopted and expanded. Cultural liaisons serve as a link between the Census Bureau and specific race/ethnic communities, explaining the objectives and introducing the enumerators and staff to local communities to give them legitimacy to be in the community and collect data. Engaging local community or neighborhood cultural liaisons informally can also open doors within ethnic enclave neighborhoods.

### **7.1.3.3 Translate census forms into more languages and consider using ACASI**

The 2010 Census results show that almost all race/ethnic groups are growing faster than non-Hispanic Whites (Humes, Jones, and Rodriguez, 2011), partly due to increasing immigration from Asia, Central and South America, and elsewhere. Our NRFU/UE observation studies in the Hispanic site, the Chinese site, and on the two American Indian reservations document problems enumerators and respondents faced when respondents could speak English well and when census forms and information sheets were not in their languages and there were not enough bilingual enumerators to conduct the interviews.

The authors recommend more research and development of forms and other materials in more foreign languages to reduce as much as possible the translation-on-the-fly approach to interviewing that has been documented. Recognition of variation in languages and dialects within some minority groups, such as Mandarin and Cantonese, and the Toisanese dialect among the Chinese in this study, for example, are important for improving measurement across population groups.

Another recommendation is to modify the Language Identification Card text for Chinese so that the choice indicates the respondents' specific language and dialect so that the appropriate bilingual interviewer can be assigned to do the interview; the 2010 Census version precluded that.

The authors suggest consideration of research and development of foreign language questionnaires and materials for some American Indian groups, in consultation with those groups. On the American Indian reservation, English is learned by many as a second language and some older people do not speak English well. Some of the interviews on these reservations were done with partial on-the-fly translations which may not produce comparable data.

Recognizing that it will not be feasible or practical to cover all languages fully in all sites, the authors suggest exploring the feasibility, appropriateness, and costs of using ACASI (Audio-Computer-Assisted Self-Interviewing) for allowing respondents to be able to take the instrument from the interviewer, put on headphones, and listen and respond to audio-recorded interviews in their own languages on the device. We heard this suggested as a possible application of ACASI at the International Conference on Methods for Surveying and Enumerating Reach Hard-to-Reach Populations in late 2012 and think this might be something to explore. At the same time, ACASI might also be considered for persons who are hard-of-hearing, so that they could turn up the volume.

#### **7.1.3.4 Identify and address challenges of enumerating with vague addresses and in remote areas**

The evaluations in San Francisco Chinatown, the American Indian reservation, the Big Island, and Kodiak Island, Alaska indicate extra challenges in enumerating areas where finding and accessing some respondents is more difficult. Additional resources, such as cell phone reimbursements, and in rural areas, four-wheel drive vehicles, may be needed to reach the remote populations in such areas to avoid undercounting them as much as possible. This appeared to be a factor in the American Indian site that might potentially have affected coverage if they were not able to get to all of the houses in remote areas of the reservation before the UE operation ended.

#### **7.1.3.5 Research on mobility and coverage**

The 2008 Dress Rehearsal and the 2010 Census were conducted during a severe economic recession with high unemployment and many foreclosures. This likely increased mobility and likely produced changes in household structure and coverage. How did these factors affect the 2010 Census? Research is suggested on the interaction of these factors and other changing social trends to identify and better understand the challenges that will face the Census Bureau as it prepares for the 2020 Census. The Census Bureau was able to recruit a much higher caliber of enumerators in 2010 than before because so many highly skilled, educated and experienced people were available to become enumerators and staff our local offices. What effects did that have on our enumeration? What conditions will we face for the 2020 Census and how do we prepare for them while also trying to contain costs and maintain quality?

#### **7.1.3.6 Possible Coverage Errors and Inconsistencies**

As mentioned in the limitations section and in the text, the EQ form used in personal visit interviews was designed to collect basic 2010 Census data. While it does include questions to flag potential erroneous enumerations (Question 7) and possible omissions (Question H1), these questions in and of themselves do not collect enough information to determine where each identified person should be counted. They serve as flags during processing of cases that may be sent for inclusion in the comprehensive 2010 Census CFU operation. One major reason that additional questions to determine residence were not added was the need for additional pages of census forms, which would significantly raise costs of paper, printing, and postage.

With the movement away from paper forms and with electronic reporting on the internet and hand-held devices, the constraints and costs of paper questionnaires are minimized. The 2010 Census Avoid Followup Experiment tested just a few additional questions on usual residence, location on April 1, 2010, and alternative address for each person to detect possible erroneous enumerations on the census form itself and was successful. It is suggested that consideration be given to reviewing these questions and adapting and testing them for use on electronic 2020 Census forms in an effort to reduce the numbers of cases requiring followup.

## **7.2 The CCM Operation**

### **7.2.1 Records check: numbers, types and reasons for possible coverage errors**

In this section, the authors described the partial validation study with CCM data that triangulated among four data sources to decide where each person should be counted in the 2010 Census and compared that to where the person was counted in the final 2010 Census dataset included in the final localized 2010 CCM/Census matched dataset. The authors focused on the outcomes for the observed housing units on Census Day—those who were correctly or apparently incorrectly counted at the observed housing units or omitted from it when it appeared they should have been counted there. Those who appeared to be correctly left off the observed housing unit rosters on April 1, 2010 and who should have been counted elsewhere were not included in this analysis.

The results indicated that 60 (6 percent) of the 953 people in the analysis sample had possible coverage error at the observed housing unit in the 2010 Census as of April 1, 2010, according to analysis with the final localized 2010 CCM/Census matched dataset. At the household level, these 60 persons were found in 41 (15 percent) of the 289 analysis sample housing units that were occupied on April 1, 2010. In some cases, the whole household was missed, while in others, just some of the persons associated with it were missed. It is important to note that the inconsistencies in possible coverage error found during the CCM records check fall into several groups. Some of the possible coverage errors were of the type where persons who appeared to be correctly interviewed in the CCM and all indications were that they should have been counted there in the 2010 Census dataset in the final localized 2010 CCM/Census matched dataset used in this study, but were not. In other cases, persons were included on the final 2010 Census rosters but not in the CCM interviews. In that situation, the authors cannot say whether they were correctly counted there or not because no additional information is available about such persons. Some of the others were errors in the CCM data collection, such as those where interviewer error was a factor. Coder/matchers coded up to three reasons for error for each of these interviews. Sources of error identified in the records check component in rank order of frequency included mobility/tenuousness, respondent confusion, interviewer error, respondent concealment/refusal, language barriers, and hidden housing units.

### **7.2.2 Reasons for possible coverage errors in ethnographer report cross-cutting themes and profiles**

The ethnographic reports can document the types of interviewer errors that occurred, how they were related to other sources of error, and identify how they may have contributed to possible

coverage error. In order to understand how these chronically undercounted groups may be differentially impacted by cultural and other factors, the discussion turns to selected ethnographic reports for the groups observed.

Like de la Puente (1993; 2004) the authors of this evaluation found households with mobility, language barriers and irregular housing (i.e., hidden housing units) as reasons for possible coverage error. We found specific evidence of these situations in the persistently undercounted groups as well as in other sites. Like the first ethnographic studies cited earlier, we saw evidence of respondent confusion and concealment as reasons for possible coverage error. In the past census ethnographies, researchers were present elsewhere in the environment while the censuses were taking place and interviewed respondents about their census experience, but those researchers did not have the ability to actually accompany enumerators or observe actual census enumerations.

Looking across the ethnographic reports and cases, the authors identified seven cross-cutting themes that could lead to coverage problems in these groups: 1) interviewer error; 2) difficulty in gaining access to respondents; 3) language barriers; 4) cultural issues; 5) mobility; 6) recall issues; and 7) respondent reactions to multiple cold-call visits from Census Bureau interviewers.

To our knowledge, interviewer error is a new reason for possible coverage error that has not been identified in previous decennial census ethnographic evaluations. The identification of this new factor was possible because this appeared to be the first census in which researchers were permitted to accompany interviewers and observe, tape, and debrief in live interviews with respondents. Types of interviewer errors that contributed to possible coverage error in this operation include: omitting or rewording questions, making assumptions (answering without waiting for the respondent to answer), not probing sufficiently, not otherwise adhering to the interview protocol, and sometimes missing key statements by the respondent due to inattention or distraction.

We saw how these four cross-cutting themes were sometimes closely intertwined with each other and with other previously identified factors. Interviewer error was observed in conjunction with respondent confusion as well as mobility. In situations where interviewers encountered hostile or non-native English-speaking respondents, ethnographers noted that their standardized interviewing behavior sometimes became worse (Sando, 2012). Though this is an untestable hypothesis in the current study, these types of situations could lead to differential undercoverage. There could be an interaction between barriers to enumeration (e.g., mobility, language barriers, hidden housing units) and how the interviewer behaves toward the respondent that actually exacerbates the problem of differential coverage. This area requires further investigation.

Ethnographers also observed interviewers having difficulty gaining access to respondents. Sometimes interviewers had difficulty accessing housing units because they were hidden. Sometimes physical access barriers were accompanied with respondent hostility, acknowledging that they were difficult to access on purpose. As was seen in the earlier 2010 Census results, very rural and remote areas had the challenge of lacking street numbers or names, making it difficult both for interviewers and respondents when they needed to refer to an address by name – sometimes leading to respondent confusion or incomplete information.

Language barriers have also been a persistent problem for the Census Bureau despite efforts to improve non-English language data collection tools. Most sites witnessed non-English language interviews, but the language barrier for completed interviews seemed most notable in the site that also demonstrated the most mobility and complexity. The combination of not having a questionnaire in the respondents' native language and the need for the respondent to report complex living situations made the language factor particularly notable on the American Indian reservations. Fortunately, in this study, all observed cases on the American Indian reservations had native language speakers as interviewers.

This evaluation, though, was also able to document wider effects of language barriers than previous decennial ethnographic studies because it documented what happened as monolingual English-speaking interviewers attempted to interview respondents with little or no English facility during live interviews. Ethnographers in the Hispanic site, the Middle Eastern site, and the Asian, primarily Chinese sites reported that contacts with some possible respondents in their target population resulted in leaving before the interview was started because the monolingual English-speaking interviewers could not find a way to communicate with the respondents. While it is likely that these non-English speaking households would be revisited by bilingual interviewers at a later time in the CCM PI operation, the need for other bilingual interviewers to go back to these housing units increases the time and costs of field enumeration. Additionally some monolingual interviewers did muddle through the interviews in English with some respondents who had trouble understanding them, when an interview in the respondents' language would have been a better choice and less wearing on both the interviewer and respondent.

Specific cultural variations were also observed within some chronically undercounted groups. In the Alaskan Native site, mobility for work was particularly notable and distinct from the other sites. Because of the remoteness, and the particular type of work that these respondents report, the living situations surrounding work life in rural Alaska pose specific challenges for surveying that make this site at risk for undercoverage.

On the two American Indian reservations, the language use and the fluidity associated with the kinship network and households could make this a difficult group to survey. These challenges also highlight the importance of hiring people who know the language and customs to conduct interviews – thus being able to understand and translate the living situation into the survey report.

The African American site demonstrated similar patterns to past ethnographic work, despite efforts to allay respondents' fears that the 2010 Census was something that could be used against them. Respondents were observed concealing information about their household members' mobility, possibly for fear that the government would use it against them.

In the Native Hawaiian site, our ethnographic study showed a disproportionate number of possible errors. The records check identified mobility as the source of most possible coverage errors in this site, but the ethnographer also noted that interviewer error contributed as well. Some interviewers' lack of cultural sensitivity in a culture with a history of distrust toward the government was shown by the ethnographer to lead to coverage mistakes.



In the Native Hawaiian site, the Hispanic site, the Middle Eastern site, and the Chinese site, a shortage or absence of interviewers of the target race or ethnicity group, as well as of interviewers who were bilingual was a factor in a number of initial contacts that did not lead to observed interviews because monolingual English-speaking interviewers could not communicate with the non-English speaking respondents. While there is no way to assess the impact of this on whether these housing units were eventually enumerated correctly or had possible coverage error, these situations did require at least one more followup interview with a bilingual interviewer, which increases the time and costs of interviewing persons with little or no facility in English.

Additionally in the Native Hawaiian site, the Hispanic site, and the Middle Eastern site, there were heightened tensions with broadcasted calls for Native Hawaiians and Middle Easterners to boycott the 2010 Census, and increased concern for Hispanics with the restrictive Arizona immigrant law and for the Middle Easterners with the boycott and with the sensitive issue that summer of Muslims planning to build a mosque near the prior Twin Towers site in New York City. These wider factors could potentially have led to higher refusal rates, but that cannot be measured in this evaluation except when potential respondents specifically mention them.

Although non-Hispanic Whites were overcounted in the 2010 Census, some of the same types of errors are documented in the non-Hispanic White site; as mentioned, four were omissions in one CCM interview, caused by an interviewer who was distracted by a computer error and failed to record four outmovers, but which the ethnographer identified in the debriefing. These were people who were going through a divorce and left around Census Day; they were not included in the 2010 Census in this housing unit or elsewhere in the final localized 2010 CCM/Census matched dataset. In that case, mobility/tenuousness is intertwined with interviewer error.

The Generalized site had a few errors. Mobility/tenuousness, respondent confusion and a possible processing error affected one case, and a hidden housing unit was a factor in another.

### **7.2.3 Research implications and recommendations**

The authors turn now to identify research implications and offer recommendations for improving enumeration and coverage across race/ethnic groups, based on this evaluation for 1) reducing interviewer error and increasing cultural awareness of minorities; 2) translating materials into more languages and conducting interviews with non-English speakers; and 3) amending the 2010 Census residence rule and situations document.

#### **7.2.3.1 Reducing Interviewer Error and Increasing Cultural Awareness of Minorities**

Frequent major changes to questions documented in the behavior coding study and in this section are a concern because they may affect the accuracy, completeness, validity, reliability and comparability of census data. The authors offer six suggestions for reducing interviewer error during interviews:

The new finding that interviewer error was linked to possible coverage error in this study suggests two potential areas of improvement. The first area involves recommending improvements to training and monitoring, while the second involves suggestions for questionnaire improvement.

The authors offer five suggestions for reducing interviewer error during interviews. First, it is suggested that a new training module be developed and tested that goes beyond just reminding interviewers to read all questions as worded. The authors suggest that the new module also explain: 1) how questions are iteratively developed and tested and 2) why it is important for the accuracy, completeness, and quality of census data for interviewers to read questions as worded; and identify 3) the types of enumeration and coverage errors that can result when questions are not read properly. If interviewers understand why it is important for them to read questions as worded, they may be more likely to read them as written than if they are just told to read verbatim without any explanation.

Second, this instruction could be supplemented by having interviewers actively participate in role playing with scripts of difficult situations, such as those with reluctant, impatient, or hostile respondents, or those that involve difficult living situations, perhaps based on some case studies in this report. Interviewers should take turns playing the roles of both the interviewer and the respondent so they get a feel for how to respond. The whole group should then discuss what worked and did not work and how they can apply those insights to improve their future enumeration.

Third, one of the ethnographers analyzed the dynamic effects of respondent reluctance, anger, and resistance on interviewer behavior over the course of the interview (Sando, 2012). This is a fruitful line of research that goes beyond behavior coding and the authors suggest more research in this area. Such research could lead to scripts and recommendations for role playing scenarios for handling such difficult situations that could be incorporated into future training.

Fourth, to avoid interviewer error associated with culturally insensitive behavior, the authors recommend that new cultural awareness and sensitivity modules be developed that identify and discuss the racial and ethnic minorities in the local interviewing area and that advise on appropriate behavior and issues that may arise. For example, in Detroit and Dearborn, it would have been useful to have had a module on Middle Easterners and Ramadan and how that might affect responses by Arabs and other non-Arab Muslims in the area. Role playing here, too, might help to prepare interviewers.

Fifth, the authors concur with Leeman and Marsh (2013) in recommending that more monitoring be done of interviewer behavior beyond their initial trainee observations. We suggest two approaches. One is to have crew leaders accompany and observe interviewers as they conduct live interviews more regularly. The other is to consider use of CARI (Computer-Assisted Audio-Recording) that can record some or all live interviews. If interviewers are told that random checks will be made of their interviews, they might be more careful in reading questions as worded.

Sixth, a few interviewers spread among some sites reworded or omitted a substantial number of questions deliberately. Some of these openly explained they did this to make the interview more conversational, others bluntly said they did it to shorten the interview and raise their productivity vis-à-vis other interviewers in the site. Some ethnographers were told by the interviewers that the more interviews completed per unit time, the more work the interviewers got and the greater their prospects for being rehired. The implication here is that the interviewer incentive structure may be a bit too much tilted toward rewarding interviewers for the quantity of completed

interviews, rather than the quality of them. Fortunately, this happened with a relative handful of interviewers across several of the field sites, but Field managers may want to consider how to address this.

Because interviewer rewording or omission of questions can sometimes indicate that questions are not clear or are difficult or too long, the authors suggest that survey researchers examine the behavior coding results in the Leeman/Marsh study and data in this evaluation to identify and target problematic questions for possible revision and testing

#### 7.2.3.2 Translations and Interviews in More Languages

Language barriers were specifically linked to a Hawaiian case of possible coverage error of three persons, but they also had a wider effect in the Hispanic site, the Middle Eastern site, the Asian, primarily Chinese, site and the Generalized site when monolingual English interviewers could not communicate with non-English speaking respondents and thus could not start or complete the interview without the help of a bilingual interviewer. There were either too few or no bilingual interviewers in some sites. Racial, ethnic, and linguistic diversity are increasing; it would be well to prepare for a greater need to conduct interviews in other languages in the 2020 Census.

The research implication in this study is that non-English –speaking respondents will need a bilingual interviewer and, if one is not available during the first interview, at least two contacts will be needed to complete the interview, unless some alternative approach to interviewing them can be found. The first suggestion would be to hire more bilingual interviewers so that interviews could be done on the first visit. Having more bilingual interviewers to conduct foreign language interviews in person or by telephone would be the best option. However, respondents speak many languages and it may not be feasible or cost effective to hire them for all languages.

Recognizing this, the authors suggest exploring the possible use of ACASI (Audio-computer-assisted self-interviewing) for allowing respondents to be able to take the instrument from the interviewer, put on headphones, and listen and respond to audio-recorded interviews in their own languages already loaded onto the device. This might reduce the need for repeat visits and for extra bilingual interviewers. We heard this suggested as a possible application of ACASI at the Hard-to-Reach Conference in 2012 and think this might be something to explore. At the same time, ACASI might also be considered for persons with impaired hearing, so that they could turn up the volume.

To support either of these options, the CCM instrument and its supporting documents – the letter, information sheet, notice of visit, and so on—should be translated into languages beyond the current English and Spanish. Asian languages should be considered: Asians now appear to be the fastest growing group in the US. It should be recognized that some race groups have two or more mutually incomprehensible foreign languages, such as Mandarin and Cantonese, so multiple versions should be considered in those cases. The supporting letter and information sheet should be available in more languages. The Language Identification Card, shown in the Asian site to not be useful for determining whether a respondent needs a bilingual Mandarin or Cantonese Chinese speaker, should be revised.

### 7.2.3.3 Recall Issues and Moving the CCM Data Collection Closer to Census Day

As discussed, some respondent confusion about whom to include in their households was related to recall decay over the four or so months that had elapsed from Census Day to the CCM Interview Day in August, 2010. This could potentially be reduced by moving the CCM PI operation as close in time to April 1, 2010 as is feasible. Perhaps the move to adaptive design and faster automated processing of cases will make this possible. A reduction in the number of overall cold-call visits to respondents across operations and quality assurance checks would reduce the frustration and resistance some respondents to multiple visits documented in several sites.

### 7.2.3.4 Amending the Residence Rule and Residence Situations Document

The ethnographer in the Alaska Native site identified cases of men living seasonally on commercial fishing vessels that are not directly covered in the key residence rule and residence situations document for the 2010 Census (Lamas, 2009). She suggested they be counted at their onshore residences, but this was based on a very small number of cases. The authors suggest that this situation be researched and that consideration be given to amending the residence rule and situations document to instruct where such persons should be counted.

The move toward using multiple new devices and internet reporting for 2020 Census data collection could potentially have implications for the 2010 Census version of the Residence Rule and Situations document (Lamas, 2009). It may be wise to consider possible implications during the early rounds of testing for the 2020 Census.

### 7.2.3.5 Race, Hispanic Origin, Household Structure and Coverage as the US Diversifies

The 2010 Census shows that the country is rapidly diversifying, with almost all racial and ethnic groups growing faster than non-Hispanic Whites (Humes, Jones, and Ramirez, 2011), partly due to increasing immigration from Asia, Central and South America, and elsewhere. Other research indicates that minorities are more likely to live in some types of complex households than non-Hispanic Whites (Schwede, 2007) and prior census results indicated that complex households increased from 18 percent in 1990 to 21 percent in Census 2000. The comparable research has not yet been done to learn if that trend continued over the decade leading to the 2010 Census. Other research shows that complex family households in general increased during the 2007-2009 years of the recession and were associated with unemployment (Elliott, Young, and Dye, 2011). Additionally, research on young adults indicated that the proportion of adult children older than college students from ages 25 to 34 living with their parents increased from 22 percent in 2010 to 29 percent a year or so later (Parker, 2012) and that those from ages 25 to 29 increased their residential dependency and may have fared worse than others from the recession (Mykyta, 2012). Some of the effects of the recession in terms of higher rates of unemployment and home foreclosures lingered through the 2010 Census and CCM PI operations.

Census Bureau projections indicate that the country will reach majority minority status sometime in the 2040s. Challenges to maintaining or improving coverage may increase in tandem with the growth of persistently hard-to-count groups and the possible growth of complex households. Additional research on the interrelationships of race, Hispanic origin, changing household structure, enumeration and coverage could prove useful in efforts to maintain quality and coverage while reducing costs in the 2020 Census.

## 8. Acknowledgments

We have many people to thank for their contributions to this large and complex evaluation through the initial planning, implementation, analysis and writing over the last several years. We first thank those who contributed by making this evaluation possible and by maintaining it through hard budget times: Frank Vitrano, Dan Weinberg, the Census Integration Group, Tommy Wright, Manuel de la Puente, and Eleanor Gerber, David Whitford, and later Peter Miller, Ron Jarmin, Kevin Deardorff, Karen Medina and Melissa Therrien. We also are grateful to those who reviewed our draft study plan and helped to improve it: Tommy Wright, Yuling Pan, Nancy Bates, David Whitford, Gia Donnalley, Geraldine Burt, Darlene Moul, Hilda Dimmock, Miriam Balutis, Larry Cahoon, Yuling Pan, Tammy Adams, Pat Cantwell, Theresa DeMaio, Elizabeth Nichols, and Jennifer Childs. We very much appreciate Debbie Fenstermaker's provision of custom runs of CCM sampling frames and Nancy Bates provision of custom Census 2000 data on hard-to-count score tracts by race/ethnicity. We extend many thanks to our colleagues in Field Division—Geraldine Burt and Darlene Moul and the Regional Census Centers for helping us get approval to accompany interviewers and observe interviews, and Denver Regional administrators Cathy Lacy, Vicki McIntire, and Partnership Specialist Amadeo Shije, as well as headquarters colleagues who advised on selecting American Indian reservations: Dee Alexander, Curtis Zunigha, and Jan Jacobs. We very much appreciate Kyra Linse's contributions to planning and setting up our CCM observations. We are very grateful to Amy Sheldon, who worked with us to get approval for our ethnographers to use secure laptops and flash drives in the field and in their homes to analyze these data.

We are also indebted to the many people who worked with us during the actual conduct of this evaluation. Foremost among these are Mandi Martinez and Ryan King, who made very substantial contributions to this evaluation, along with Matt Clifton, who provided much-needed help at the beginning. We are very grateful to our 16 contract ethnographers for conducting the field studies, transcribing and summarizing their interviews and preparing and revising their detailed coverage tables and their final site reports. Our NRFU ethnographers included: Robin Albee, Rae Lesser Blumberg, Alexis Bunten, Victor Garcia, Heather McMillen-Wolfe, Robin Jarrett, Sylvester Lahren, and Dawn Lee Tu. Our CCM contract ethnographers included: Ephrosine Daniggelis, Rachel Donkersloot, Theodore Fortier, Angelina Foster, Eleanor Hubbard, Karyn Lacy, C. Alison Newby, and Ruth Sando. We extend heartfelt thanks to our coder/matchers: Mandi Martinez and Tammy Gaich who coded both census and CCM interviews; Denise Gordon, Crystal Miller, Elda Robinson, Irvin Vann, Julia Shaw, Eric Cheevers, Laura Becht, and Christina Cooper for coding/matching CCM cases; and John Jones and Kopen Henderson for their work in coding/matching NRFU cases. We very much appreciate the willingness of Hilda Dimmock and Jay Occhiogrosso in sharing staff with us part time to do coding and matching. In addition, we are indebted to Gia Donnalley for providing personnel to

train and support our staff in using PerMaRCS, for workstations, for part-time coding staff, and for continued support and encouragement throughout the evaluation. We also thank Ann Dimler for guidance on managing contracts and Carol Gentry and Jennifer Abbott for helping us with contracting. We are very appreciative of the work of Joseph Butler and Eric Conner in LTSO who prepared special laptops for the ethnographers and answered their many questions and Chad Russell who advised us on IT issues. Last, but not least, we thank all of the Regional and Local 2010 Census Office employees and field staff who made the arrangements for the observations and the 122 enumerators/interviewers who agreed to be accompanied, as well as our respondents.

Finally, during the later writing phase, we very much appreciate Jennifer Child's review of an earlier draft and her willingness to collaborate with us on a related publication (Schwede, Terry and Childs, forthcoming) as well as to include her historical review of Census Bureau ethnographic studies in the appendix. We also thank Michelle Smirnova for her contributions to that literature review. We thank our final critical reviewers for their comments to improve this report: Peter Miller, Yuling Pan, Jennifer Childs, Elizabeth Nichols, Nancy Bates, Magda Ramos, Gia Donnalley, Kyra Linse, John Jones, Patricia Sanchez, Sarah Heimel, Bryn Johnson, and Joni Thomson. We also very much appreciate our eagle-eyed editors and reviewers, including Sherri Norris, Ann Tozzi, Melissa Therrien, and Kevin Deardorff, and perhaps others. Thanks also to Joyce Fields and Donna Souders for their help and guidance related to CPEX over the years. We very grateful to Gia Donnalley, Hilda Dimmock, Miriam Balutis, and Regional Office colleagues for reviewing the ethnographer reports, especially Gia, who reviewed all of them for confidentiality issues. We also thank Peter Miller, Yuling Pan, Terry DeMaio, Kevin Deardorff, and Jim Dinwiddie for comments on a related presentation and publication.

## 9. References in the Text

Albee, R. (2011), "Ethnographic Study of 2010 Nonresponse Followup Enumeration Methods and Coverage in Non-Hispanic White Population in Independence, Missouri." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: the 2010 Census Nonresponse Followup Operation. Suitland, MD: U.S. Census Bureau. Under internal review for possible future web posting in the Center for Survey Measurement Research Report Series.

Aschenbrenner, J. (1975), *Lifelines: Black Families in Chicago*. New York: Holt, Rinehart and Winston.

Bates, N. and E. Gerber. (1998), "Temporary Mobility and Reporting of Usual Residence." *Survey Methodology*, 24(1): 89-98.

Bates, N. and M. Mulry. (2011), "Using a Geographic Segmentation to Understand, Predict, and Plan for Census and Survey Mail Nonresponse." *Journal of Official Statistics*, December.

Bell, W. (2004), "Distant relatives have lower census inclusion probabilities than immediate family." Email message to David Whitford: Chief of Decennial Statistical Studies Division. May 27, 2004.

Blumberg, R. (2006), "Gender, Economy, and Kinship in Complex Households among Six U.S. Ethnic Groups. Who Benefits? Whose Kin? Who Cares?" In L. Schwede, R.L. Blumberg, and A.Y. Chan (editors), *Complex Ethnic Households in America*, MD: Rowman & Littlefield Publishing, Inc.

Blumberg, R. (2011), "Comparative Ethnographic Study of Enumeration Methods and Coverage: The Generalized Site in Broward County, Florida in the 210 Census Nonresponse Followup Operations." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: the 2010 Census Nonresponse Followup Operation. Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. Suitland, MD: U.S. Census Bureau.

Brandon, P. (2008), "New Approaches to Household Diversity and Change: Guest Editors' Note." *Journal of Population Research* Special Issue: New Approaches to Household Diversity and Change, 25(3).

Brown, S. and W. Manning. (2011), *Counting Couples, Counting Families, Full Report*. Counting Couples, Counting Families Conference, July 19-20 at the National Institutes of Health in Bethesda, MD. Bowling Green, OH: National Center for Family and Marriage Research, Bowling Green University.

Bunten, A. (2011), "Comparative Ethnographic Studies of Enumeration Methods and Coverage in Race/Ethnic Groups: Alaska Natives and Others in the Kodiak Borough 2010 Nonresponse Followup Operation." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: the 2010 Census Nonresponse Followup Operation. Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Childs, J. (2006), "Not the Typical Household: Rural Whites in New York." In L. Schwede, R.L. Blumberg, and A.Y. Chan (editors), *Complex Ethnic Households in America*, MD: Rowman & Littlefield Publishing, Inc.

Childs, J. H. and N. Jurgenson. (2011), "Behavior Coding of the 2010 Census Nonresponse Followup (NRFU) Interview Report." 2010 Census Program for Evaluations and Experiments. 2010 Census Planning Memoranda Series Memo #150. August 30, 2011. Suitland, MD: U.S. Census Bureau.

Compton, E. and Bentley, M. (2012), "2010 Census Nonresponse Followup (NRFU) Contact Strategy Experiment Report." 2010 Census Program for Evaluations and Experiments. 2010 Census Planning Memoranda Series No. 174. February 17, 2012. Suitland, MD: U.S. Census Bureau.

Craver, A. (2006), "Household Adaptive Strategies among the Inupiat." In L. Schwede, R.L. Blumberg, and A.Y. Chan (editors), *Complex Ethnic Households in America*, MD: Rowman & Littlefield Publishing, Inc.

Daniggelis, E. (2011), "A Comparative Ethnographic Study of Enumeration Methods and Coverage in a Native Hawaiian and Pacific Islander site in the 2010 Census Coverage Measurement Survey Operation." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: the 2010 Census Coverage Measurement Survey Person Interview Operation. Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

de la Puente, M. (1993), "Why are People Missed or Erroneously Enumerated in the Census: A Summary of Findings from Ethnographic Research." In *Proceedings of the 1993 Research Conference on Undercounted Ethnic Populations*, pp. 29-66.

de la Puente, M. (2004), "Census 2000 Ethnographic Studies. Census Topic Report Number 15." *Census 2000 Testing, Experimentation and Evaluation Program TR-15*. Issued March, 2004.

de la Puente, M. and Stemper, D. (2004), "The Enumeration of Colonias in Census 2000: Perspectives of Ethnographers and Census Enumerators." Census 2000 Evaluation J.4: Ethnographic Studies. September 22, 2003. Suitland, MD: U.S. Census Bureau.  
Available at: <http://www.census.gov/pred/www/rpts/TR15.pdf>

Donkersloot, R. (2011), "Comparative Ethnographic Studies of Enumeration Methods and Coverage in Race/Ethnic Groups in the Census Coverage Measurement Person Interview Operation: Kodiak Island Borough, Alaska." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: the 2010 Census Coverage Measurement Survey Person Interview Operation. Suitland, MD: U.S. Census Bureau.  
Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Elliott, D. R. Young, and J. Dye. (2011), "Variation in the Formation of Complex Family Households during the Recession." Symposium Paper presented at the

National Council on Family Relations Annual Conference, Orlando, Florida. November 16-19, 2011. Suitland, MD: U.S. Census Bureau. SEHSD Working Paper Number 2011-32.

Ellis, Y. (1994), "Categorical Data Analysis of Census Omissions." Decennial Statistical Studies Division 1990 REX Memorandum Series #PP-10. July 26. Suitland, MD: U.S. Census Bureau.

Fallica, H., S. Heimel, G. Jackson, and B. Zhang. (2012), "2010 Census Update Enumerate Operations Assessment Report." Census Program for Evaluations and Experiments. 2010 Census Planning Memoranda Series No. 245, October 11, 2012. Suitland, MD: U.S. Census Bureau.

Fortier, T. (2012), "Comparative Ethnographic Studies of Enumeration Methods and Coverage in Race/Ethnic Groups in the Census Coverage Measurement Person Interview Operation: American Indian Sites." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: the 2010 Census Coverage Measurement



Survey Person Interview Operation. Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Foster, A. (2012), "Comparative Ethnographic Studies of Enumeration Methods and Coverage in Race/Ethnic Groups in the Census Coverage Measurement Survey: The Arab and Middle Eastern Site." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: the 2010 Census Coverage Measurement Survey Person Interview Operation. Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Garcia, V. (1992), "Results from an Alternative Enumeration in a Mexican and Mexican American Farm Worker Community in California: Ethnographic Evaluation of the Behavioral Causes of Undercount." Final Coverage Report for Joint Statistical Agreement 89-29. Washington, D.C.: Center for Survey Methods Research, Bureau of the Census, Department of Commerce.

Garcia, V. (1995), "Finding and Enumerating Migrants in Mexican Enclaves of the U.S. Northeast: The Case of Southern Chester County, Pennsylvania." Washington, D.C.: Center for Survey Methods Research, Bureau of the Census, Department of Commerce.

Garcia, V. (2011), "2010 Census CPEX: Ethnographic Evaluation of Enumeration Methods and Coverage: Nonresponse Followup 2010 Observations of Hispanics and Others in the Greater Dallas Area." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report. Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Garcia, V. (2012), "Enumerating Hispanic Immigrants: Observations from Census Studies and Recommendations for an Accurate Count." Paper presented at the International Conference on Methods for Surveying and Enumerating Hard-to-Reach Populations, November 3, 2012 in New Orleans, Louisiana.

Garcia, V., Lopez-Ariza, B., and Marinez, J. (2008), "Exploring Undercounts in the Agriculture Census: An Alternative Enumeration of Hispanic Farmers in Southwestern Michigan." *Journal of Latino-Latin American Studies (JOLLAS)*, 3(2), 47-58.

Gerber, E. (1994), "Respondents' Understanding of Residence Terminology in Cognitive Research and in the Living Situation Survey." *Proceedings of the Bureau of the Census 1994 Annual Research Conference*, March 20-23, 434-49. Suitland, MD: U.S. Census Bureau.

Gerber, E. (2003), "Privacy Schemas and Data Collection: An Ethnographic Account." Census 2000 Testing, Experimentation and Evaluation Ethnographic Report. U.S. Census Bureau. Suitland, MD: U.S. Census Bureau.

<http://www.census.gov/pred/www/rpts/Privacy%20Schemas%20Final%20Report.pdf>

Goerman, P. (2006), "Making Ends Meet: The Complex Household as a Temporary Survival Strategy among New Latino Immigrants to Virginia," in L. Schwede, R.L. Blumberg, and A.Y. Chan. Lanham (eds.), *Complex Ethnic Households in America*, MD: Rowman & Littlefield Publishing.

Hamid, A. (1992), "Ethnographic Follow-up of a Predominately African American Population in a Sample Area in Central Harlem, New York City: Behavioral Causes of the Undercount of the 1990 Census." *Ethnographic Evaluation of the 1990 Decennial Census, Report #11*. Washington, DC: Bureau of the Census.

Heimel, S. and R. King. (2012), "2010 Census Effectiveness of Unduplication Evaluation, 2010." Census Program of Evaluations and Experiments Report. 2010 Census Program for Evaluations and Experiments. 2010 Census Planning Memoranda Series No 244, October 10, 2012. Suitland, MD: U.S. Census Bureau.

Hogan, H. (1993), "The 1990 Post-Enumeration Survey: Operations and Results." *Journal of the American Statistical Association*, 88(423):1047-1060.

Holmes, B. (2003), "African American Households in Transition: Structure, Culture, and Self-definition. In L. Schwede, R.L. Blumberg, and A.Y. Chan. Lanham (eds.), *Complex Ethnic Households in America*, MD: Rowman & Littlefield Publishing.

Hubbard, E. (2011), "'I'm Not White or Irish, I'm American': A Comparative Ethnographic Study of Enumeration Methods and Coverage in a Predominately White Setting in the 2010 Census Coverage Measurement Survey." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: the 2010 Census Coverage Measurement Survey Person Interview Operation. Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Humes, K.R., Jones, N.A., and Ramirez, R. R. (2011), *Overview of Race and Hispanic Origin: Census 2010 Brief*. Washington, DC: Economics and Statistics Administration, U.S. Census Bureau. Available online at <http://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf>.

Hunter (Childs), J. de la Puente, M., and Salo, M. (2003), "Comparative Ethnographic Research on Mobile Populations." Census 2000 Evaluation J.3: Ethnographic Studies. July 17, 2003. Suitland, MD: U.S. Census Bureau.

Jackson, G. and K Wechter (2012), "2010 Census Avoid Followup Evaluation Report. 2010." Census Program for Evaluations and Experiments. 2010 Census Planning Memoranda Series No. 236, September 7, 2012. Suitland, MD: U.S. Census Bureau.

Jarrett, R. (2010), "Comparative Ethnographic Study of Enumeration Methods and Coverage in Race/Ethnic Groups: African Americans and Others in the 2010 Census Nonresponse Followup Operation." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report. Suitland, MD: U.S. Census Bureau. Under review for possible

future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Kang, T. (2006), "Household Structure and its Social and Economic Functions among Korean American Immigrants in Queens, New York: An Ethnographic Study" in L. Schwede, R.L. Blumberg, and A.Y. Chan.(eds.), *Complex Ethnic Households in America*, MD: Rowman & Littlefield Publishing, Inc.

Lacy, K. (2011), "An Ethnographic Study of Enumeration Methods and Coverage in a 2010 Census Coverage Measurement Survey Operation Primarily among Blacks in a Chicago Field Site." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: the 2010 Census Coverage Measurement Survey Person Interview Operation. Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Lahren, S. (2012), "Comparative Study of Enumeration Methods and Coverage on a Rural American Indian Reservation: A Case Study. Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: The 2010 Census Update Enumerate Operation." Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Lamas, E. (2009), "Residence Rule and Residence Situations for the 2010 Census." Memorandum from Enrique Lamas, Chief, Population Division, to Frank Vitrano, Chief, Decennial Management Division, April 24, 2009. Suitland, MD: U.S. Census Bureau.

Lee Tu, D. (2011), "Comparative Ethnographic Study of Enumeration Methods and Coverage in a Chinese Site during the 2010 Census Nonresponse Followup (NRFU) Operation." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: the 2010 Census Nonresponse Followup Operation. Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Leeman, J. and H. Marsh (2013), "Behavior Coding Report of 2010 Census Coverage Measurement Person Interviews." 2010 Census Program for Evaluations and Experiments. 2010 Census Planning Memoranda Series Memo #251. Washington, DC: U.S. Census Bureau.

Letourneau, E. (2012), "Mail Response/Return Rates Assessment." 2010 Census Program for Evaluations and Experiments. 2010 Census Planning Memoranda Series, Memo #198, June 6, 2010. Suitland, MD: U.S. Census Bureau.

Martin, E. (2007), "Strength of Attachment: Survey Coverage of People with Tenuous Ties to Residences," in *Demography* 44(2): 427-440.

Martin, E., Brownrigg, L., and Fay, R. (1990), "Results of 1988 Ethnographic Studies of Census Coverage and Plans for 1990." Paper presented to the Census Advisory Committees of the

American Statistical Association and then Population Statistics at the Joint Advisory Committee Meeting, Alexandria, VA.

McMillen-Wolfe, H. (2011), "Ethnographic Evaluation of Enumeration Methods and Coverage in the 2010 Nonresponse Followup Operation on the Island of Hawaii, with Special Focus on Native Hawaiians and Other Pacific Islanders." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: the 2010 Census Nonresponse Followup Operation. Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Mule, T. (2012), "Census Coverage Measurement Estimation Report: Summary of Estimates of Coverage for Persons in the United States." DSSD 2010 Census Coverage Measurement Memorandum series #10`0-G\_01. Washington, DC: U.S. Census Bureau.

Mule, T. (2003), "A.C.E. Revision II Results: Change in Estimated Net Undercount." Memorandum from Dawn Haines to Donna Kostanich, March 4, 2003. Suitland, MD: U.S. Census Bureau.

Mykyta, L. (2012), "Economic Downturns and the Failure to Launch: The Living Arrangements of Young Adults in the U.S. 1995-2011." SEHSD Working Paper 2012-24. Washington, DC: US Census Bureau.

National Research Council. (2004), *The 2000 Census: Counting under Adversity*. Panel to Review the 2000 Census, p. 229, Table 6.7. Constance F. Citro, Daniel L. Cork, and Janet L. Norwood, eds. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Newby, C. A. (forthcoming), "Comparative Ethnographic Studies of Enumeration Methods and Coverage in Race/Ethnic Groups in the Hispanic Site during the 2010 Census Coverage Measurement Survey Person Interview Operation." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: the 2010 Census Coverage Measurement Survey Person Interview Operation. Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Nichols, E., and Childs, J.H. (2007), "Respondent Debriefings Conducted by Experts: A New Qualitative Methods for Questionnaire Evaluation." U.S. Census Bureau Report RSM2007/39. Suitland, MD: U.S. Census Bureau.

Available at: <http://www.census.gov/srd/papers/pdf/rsm2007-39.pdf>

Nichols, E., and Childs, J.H. (2009), "Respondent Debriefings Conducted by Experts: A Technique for Questionnaire Evaluation." *Field Methods*, 21(2) 115-132.

Nichols, E., J. H. Childs, L. Schwede, J. Bibb, V. Smith, S. Norton, J. Burnham, T. Adams, E. Martin. (2007), "2006 Census Coverage Measurement Person Interview Trip Reports: July, 2006." DSSD 2010 Census Coverage Measurement Series #2006-D7-11.

O'Hare, W. (2013), "Data on the Undercount of Young Children in the U.S. Decennial Census." Center for Statistical Research and Methodology Seminar. January 16, 2013. Suitland, MD: U.S. Census Bureau.

Olson, D. and R. Sands. (2012), "Census Coverage Measurement Report: Net Coverage Comparisons with Post-Stratification." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-G-12. Suitland, MD: U.S. Census Bureau.

Pan, Y. and S. Lubkemann (2012), "Observing Census Enumeration of Non-English Speaking Households in the 2010 Census." 2010 Census Program for Evaluations and Experiments. 2010 Census Planning Memoranda Series, Memo #249. Suitland, MD: U.S. Census Bureau.

Parker, K. (2012), "The Boomerang Generation: Feeling Okay about Living with Mom and Dad." Pew Research Center Report. Released March 15, 2012.

Robinson, G., A. Bruce, G. Mills and E. Love. (2008), "The Planning Database." August, 2010. Suitland, MD: U.S. Census Bureau.

At: [www.census.gov/cic/docs/2008\\_08-2010\\_Planning\\_Databse-Erin\\_Love.ppt](http://www.census.gov/cic/docs/2008_08-2010_Planning_Databse-Erin_Love.ppt).

Sando, R. (2011), "Comparative Ethnographic Study of Enumeration Methods and Coverage in a Generalized Population Site: 2010 Census Coverage Measurement Operation." Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups Site Report: the 2010 Census Coverage Measurement Survey Person Interview Operation. Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Sando, R. (2012), "Exploring the Effects of Respondent Behavior on the interview Process and Data Quality in a 2010 Census Case Study Evaluation." Presented at the International Conference on Methods for Surveying and Enumerating the Hard-to-Reach Populations in New Orleans. November 1, 2012. Suitland, MD: U.S. Census Bureau. Under review for possible future web posting in the Center for Survey Measurement Research Report Series. U.S. Census Bureau.

Schwede, L. (2003), "Complex Households and Relationships in the Decennial Studies and in Ethnographic Studies of Six Race/Ethnic Groups." Census 2000 Testing, Experimentation and Evaluation Program Ethnographic Research Report. Suitland, MD: U.S. Census Bureau.  
<http://www.census.gov/pred/www/rpts/Complex%20Households%20Final%20Report.pdf>

Schwede, L. (2004), "Household Types and Relationships in Six Race/Ethnic Groups: Conceptual and Methodological Issues." Paper presented at the American Association for Public Opinion Research Annual Meetings. Published in *2005 Joint Statistical Meetings Proceedings, Section on Survey Research Methods*, pp. 4991-4998. Available at:  
<http://www.amstat.org/sections/SRMS/Proceedings> in AAPOR paper section.

Schwede, L. (2006), "Who Lives Here? Complex Ethnic Households in America." In *Complex Ethnic Households in America*, edited by L.Schwede, R. Blumberg, and A. Chan. Lanham, MD: Rowman & Littlefield Publishers, Inc., pages 280-320.

Schwede, L. (2007), "A New Focus: Studying Linkages among Household Structure, Race/Ethnicity, and Geographical Levels, with Implications for Census Coverage." SRD Report RSM 2007/38. Suitland, MD: U.S. Census Bureau. Abstract available at: <http://www.census.gov/srd/www/abstract/rsm2007-38.html>  
Paper available at: <http://www.census.gov/srd/papers/pdf/rsm2007-38.pdf>

Schwede, L. (2008a), "Using Multiple Data Sources to Identify Types and Sources of Coverage Error on an Indian Reservation." Presented at the JSM Annual Meetings in Denver. *2008 JSM Proceedings, Section on Survey Research Methods*, pp. 2485-2492, available at: <http://www.amstat.org/sections/srms/proceedings/y2008/Files/301574.pdf>

Schwede, L. (2008b), "Changes in the Distribution of Complex and Noncomplex Households in the United States: 1990 and 2000 Censuses." Unpublished research typology for classifying household types in this evaluation. Suitland, MD: U.S. Census Bureau.

Schwede, L. (2009), "What Can We Learn from Within-Site Pretesting of the Census 2010 Enumerator Questionnaire on the Navajo Reservation?" Paper presented at the Society for Applied Anthropology Annual Meetings in Santa Fe in March.

Schwede, L. (2011a), "Comparative Ethnographic Studies of Enumeration Methods and Coverage in Race/Ethnic Groups." 2010 Census Evaluations, Experiments and Assessments Study Plan. Suitland, MD: U.S. Census Bureau. Available within the Census Bureau on the DMD Portal.  
<http://dv102.decennial.census.gov:7778/portal/page/portal/2010censusplanninghome/dmd2010prg0000001/DMD2010MEM0000001/DMD2010MEM0000003/2010%20Census%20Planning%20Memo%20135.pdf>

Schwede, L. (2011b), "Methodological Challenges in Implementing Comparative Ethnographic Evaluation Research in Race/Ethnic Populations in the 2010 Census: an Asian Case Study." Presented at the AAPOR Annual Meetings in Phoenix on May 12, 2011. *Proceedings of the 2011 Joint Statistical Meeting Section of Survey Research Methods*.

Schwede, L. (2012), "Measuring Relationships and Complex Household Types across Race/Ethnic Groups in the Census as the U.S. Diversifies." Invited presentation to the Interagency Working Group on Measuring Relationships in Federal Surveys in Washington, D.C. May 31, 2012. Suitland, MD: U.S. Census Bureau.

Schwede, L. (forthcoming), "Comparative Ethnographic Studies of Enumeration Methods and Coverage in Race/Ethnic Groups Site Report: the 2010 Census Coverage Measurement Survey Person Interview Operation in the Asian, Primarily Chinese, Site." Suitland, MD: U.S. Census Bureau.

Schwede, L., R.L Blumberg, and A.Y. Chan. (Eds.) (2006), *Complex Ethnic Households in America*. Lanham, MD: Rowman & Littlefield Publishing.

Schwede, L., Sorokin, A. and Yelei, V. Wake. (2011), “Cognitive Testing of the Census 2010 Experimental Overcount Questions for the Census 2010 Alternative Questionnaire Experiment: Final Project Report.” Suitland, MD: U.S. Census Bureau. Research Reports (Survey Methodology #2011-02). <http://www.census.gov/srd/papers/pdf/rsm2011-02.pdf>

Schwede, L. and R Terry. (2012a), “How Beneficial Was Language Fluency of Observers in a Comparative Field-based Evaluation of the 2010 Census across Race/Ethnic Groups?” Paper presented at the Comparative Survey Design and Implementation (CSDI) Annual Conference on March 22, 2012 in Washington, D.C.

Schwede, L. and R. Terry. (2012b), “Guidelines for Matching Ethnographers to Targeted Race/Ethnic Sites in Cross-Cultural Survey Evaluations: How Well Did They Work in a 2010 Census Evaluation?” Paper presented at the American Association for Public Opinion Research Conference in Orlando, Florida on May 17.

Schwede, L., R. Terry, and J.H. Childs. (2012), “Empirical Ethnographic Evaluations of Coverage of Hard-to-Count Race/Ethnic Minorities in Decennial Censuses in the United States.” Presentation at the International Conference on Methods for Surveying and Enumerating Hard-to-Reach Populations. New Orleans, November 1, 2012.

Schwede, L., R. Terry, and J.H. Childs, (forthcoming), “Ethnographic Evaluations on Coverage of Hard-to-Count Minorities in U.S. Decennial Censuses,” in R. Tourangeau, N. Bates, B. Edwards, T. Johnson, and K.M. Wolter (eds.), *Interviewing Hard-to Survey Population*. Cambridge University Press.

Terry, R., J. H. Childs, L. Schwede, and R. King. (2012), “Using Behavior Coding to Examine the Role of Interviewer and Respondent Race/Ethnicity When Enumerating Hard-to-Reach Racial Minorities.” Paper presented at the International Conference on Methods for Surveying and Enumerating Hard-to-Reach Populations. New Orleans, November 1, 2012.

Tongue, N. (2006), “I Live Here and I Stay There: Navajo Perceptions of Households on the Reservation”. In L. Schwede, R.L. Blumberg, and A.Y. Chan (eds.), *Complex Ethnic Households in America*, MD: Rowman & Littlefield Publishing Company.

Sun Sentinel. (2010), “Broward has Florida’s Top Foreclosure Rate,” 11 February, 2010.

U.S. Census Bureau (2012), 2012 Census Coverage Measurement Results News Conference.

May 22, 2012. Available at:

[http://2010.census.gov/news/pdf/20120512\\_ccm\\_newsconf\\_slides.pdf](http://2010.census.gov/news/pdf/20120512_ccm_newsconf_slides.pdf)

U.S. Census Bureau (2003), “Technical Assessment of A.C.E. Revision II, Table 1.” Washington, D.C: U.S. Census Bureau. March 12, 2003.

U.S. Census Bureau (2009a), “2010 Census Nonresponse Followup (NRFU) Enumerator Manual,” D-547, July 2009. Economic and Statistics Administration, U.S. Department of Commerce, Washington, D.C.

U.S. Census Bureau. (2009b), “Update Enumerate Enumerator Manual.” D-1202, September, 2009. Economic and Statistics Administration, U.S. Department of Commerce, Washington, D.C.

U.S. Census Bureau. (2010), “Race and Hispanic Origin of the Foreign-Born Population in the United States: 2007.” Suitland, MD: U.S. Census Bureau. Available at: <http://www.census.gov/prod/2010pubs/acs-11.pdf>

U. S. Census Bureau. (2012a), “2010 Census Coverage Measurement Results News Conference: Transcript.” May 22, 2012. Available at: [http://www.census.gov/2010census/news/pdf/20120522\\_ccm\\_newsconference\\_transcript.pdf](http://www.census.gov/2010census/news/pdf/20120522_ccm_newsconference_transcript.pdf)

U.S. Census Bureau. (2012b), “Census Coverage Measurement Results News Conference: Slides.” May 22, 2012. [http://www.census.gov/2010census/news/pdf/20120512\\_ccm\\_newsconf\\_slides.pdf](http://www.census.gov/2010census/news/pdf/20120512_ccm_newsconf_slides.pdf)

Valentine, C.B and Valentine B.L. (1971), “Missing Men: A Comparative Methodological Study of Underenumeration and Related Problems.” Report prepared under Joint Statistical Agreement with the Brookdale Hospital Center, Washington. Suitland, MD: U.S. Census Bureau.

Walker, S., S. Winder, G. Jackson, and S. Heimel. (2012). “The 2010 Census Nonresponse Followup Operations Assessment.” Census Program for Evaluations and Experiments. 2010 Census Planning Memoranda Series No. 190, April 30, 2012. Suitland, MD: U.S. Census Bureau.

Zinzius, B. (2005), *Chinese America: Stereotype and Reality: History, Present, and Future of the Chinese Americans*. New York, New York: Peter Lang Publishing, Inc

.



## **Appendix A: A Brief Review of Coverage, Ethnographic Studies, and Changing Census Bureau Operations since the 1970s**

Jennifer Hunter Childs

"While the overall coverage of the [2010] census was exemplary, the traditional hard-to-count groups, like renters, were counted less well. Because ethnic and racial minorities disproportionately live in hard-to-count circumstances, they too were undercounted relative to the majority population." U.S. Census Director Robert Groves.<sup>1</sup>

"More work remains to address persistent causes of undercounting, such as poverty, mobility, language isolation, low levels of education, and general awareness of the survey." Deputy U.S. Commerce Secretary Rebecca Blank.<sup>2</sup>

In the 2010 Census in the United States, the Black, Hispanic and American Indian and Alaskan Native populations living on reservations were undercounted, while the non-Hispanic White population was overcounted (Mule, 2012). Historically, Blacks have been undercounted relative to non-Blacks and this has been empirically documented since the 1940s (Robinson, Ahmed, Das Gupta, and Woodrow, 1993; U.S. Census Bureau, 2010; Clark and Moul, 2004). This follows literally decades of work at the Census Bureau to understand the conditions surrounding, and to implement operations that will effectively be able to enumerate, hard-to-count groups in censuses. Since the late 1960s, the Census Bureau has employed ethnographic methods to study living situations that could lead to differential undercounts of the minority populations in order to improve operations to count the hard-to-count in future censuses.

The first ethnographic study was of low-income Blacks and Hispanics in Brooklyn (Valentine and Valentine, 1971). This qualitative study provided the first evidence of deliberate concealment as a reason for undercoverage. Other studies since then have added to the knowledge about motivational barriers to enumeration, which include providing "official" reports to the government, fear of prosecution, and doubts about the confidentiality of the data collection and the intent of its use (Hainer, 1987; Anderson, 1990; Bourgois, 1990). In addition to motivated misrepresentation, misunderstanding of the rules for census reporting also contributes to undercounting (Aschenbrenner, 1975; Hainer, 1987; Hudgins, Holmes, and Locke, 1990; Gerber, 1990).

As a part of the 1990 Census, several operations were implemented or improved to increase coverage and specifically improve the differential undercount (U.S. Census Bureau, 1993). The most notable of these with regard to differential coverage were the Parolee/Probationer Coverage Improvement Program and the Followup Program, the Shelter and Street Night Enumeration, the Transient Night Enumerations, and the "Were You Counted" program. These were attempts to distribute forms and/or enumerate known undercounted populations.

---

<sup>1</sup> [http://www.census.gov/newsroom/releases/archives/2010\\_census/cb12-95.html](http://www.census.gov/newsroom/releases/archives/2010_census/cb12-95.html)

<sup>2</sup> <http://www.commerce.gov/blog/2012/05/23/2010-census-time-under-budget-and-extremely-accurate>

In conjunction with the 1990 Census, a large-scale ethnography was conducted in 29 sites throughout the continental U.S. and Puerto Rico designed to understand living situations that led to a differential undercount by racial and ethnic groups (Brownrigg and de la Puente, 1992; 1993; de la Puente, 1991). Those studies identified many themes that ran through the different groups examined. Mobility, complex households, concealment and distrust, difficult-to-access (or irregular) units and language barriers were found across all of the studied groups. Mobility came in many different forms across the groups – children staying with other relatives, temporary household members due to recent immigration, homelessness, flexible household boundaries, and transient young men (de la Puente, 1993). Complex households were also evidenced in many forms across groups – overcrowded apartments, which sometimes included temporary, unrelated new immigrants, and an unclear determination of who were visitors and who should be counted as usual residents.

Motivational factors for undercounting also persevered across ethnic and race groups. The Hispanic and Asian studies showed that landlord and welfare regulations about how many people could reside in a unit and presence of illegal immigrants were related to how respondents reported in the census. Crime in unsafe neighborhoods also led to respondents to be distrustful of strangers and unlikely to talk to anyone. Illegal activity in the neighborhoods also made gleaning accurate census information more difficult. Hamid (1992) reported that “enumerator fear” resulted in undercounting Blacks in Harlem due to interviewers being afraid to enter some of the buildings. In the Hispanic site and on the American Indian reservations, ethnographers cited cultural resistance to compliance with the government as a problem for enumeration. Recent immigrants may have no particular loyalty to the United States, and if they plan to return to their country of origin, they may not feel a need to comply (Rodriguez and Hagan, 1991). For American Indians, the cultural distrust may be more deeply seated and may cause respondents to doubt the motives of the government (Bonvillain, 1991; Lobo, 1990).

Irregular housing was a large factor cited in these studies for the omission of people in typically undercounted racial and ethnic groups. A common theme was structures converted to accommodate more families or more people: illegal conversions of apartments; single-family homes or duplexes converted into multi-units; motels or hotels converted or functioning as apartment buildings; units behind or above businesses; and seemingly abandoned buildings that were actually occupied (de la Puente, 1993). The very rural nature of American Indian reservations posed additional difficulties locating housing units because of poor maps and people living in unconventional structures, such as abandoned vehicles (Moore, 1992). The final “structural,” in some senses, barrier to enumeration cited among these groups was a lack of knowledge of English, and in some cases illiteracy in the respondents’ native language as well (de la Puente, 1993b).

In Census 2000, further improvements were made to operations with the goal of reducing the historical undercount (Clark and Moul, 2004). In addition to special operations that targeted transient populations, for the first time there was paid advertising as a part of the outreach and promotions to inform the public about the census to increase response rates, both overall and in historically undercounted and hard-to-reach communities. Census 2000 was the first to document a net overcount; however, there was still a differential undercount of the Black population (Clark and Moul, 2003; Robinson and Adlakha, 2002).

Building on past success with the ethnographic methods to better understand issues of undercoverage with hard-to-count populations, six separate ethnographic studies were mounted for Census 2000 focusing on privacy, Generation X, mobile populations, complex households, social network tracing and enumeration barriers in colonias (de la Puente, 2004). Results of the studies most relevant to racial and ethnic minorities showed that many of the same themes discovered in prior ethnographies were supported. The studies on mobile populations examining youth gangs, urban American Indians, Irish Travelers and snowbirds and the studies on the colonias (border towns between the U.S. and Mexico) found enumeration barriers of mobility, irregular housing, complex and fluid households and distrust and disinterest in the process of the census, very similar to the 1990 studies (Hunter, de la Puente, Salo, 2003; de la Puente and Stemper, 2003). In addition to these, the colonias study identified issues of limited English proficiency (de la Puente and Stemper, 2003).

The ethnographic study on complex households was a series of six studies with different racial and ethnic groups, which also found enumeration problems due to mobility, but it expanded on that understanding by illuminating how differences in respondents' conceptual definitions of "household" and "usual residence" can result in enumeration errors (Schwede, 2003; Schwede, Blumberg, and Chan, 2006). In addition, the complex household study found fear, mistrust and doubt to be motivational drivers of omissions from the census. In addition to these major findings, most other barriers to enumeration from the 1990 studies were supported again in the 2000 study (Schwede, 2003; Schwede, Blumberg, and Chan, 2006).

The study on social network tracing provided additional support and clarity to initial 1990 findings (Brownrigg, 2003). This study examined characteristics of mobility and how those characteristics were related to census omissions. Like in the 1990 work, unconventional housing units were a key source of missed households (Brownrigg, 2003).

In the 2010 Census, more improvements were made to the address listing procedures to improve housing unit coverage, procedures used to resolve uncertain coverage situations were refined and a new operation for people living in transitory locations was implemented to reduce problems with coverage. Yet, the net result was still differential undercoverage of the Black, Hispanic and American Indian and Alaskan Native populations living on reservations.

While the Census Bureau has been made aware of these barriers to enumeration through ethnographies throughout these decades and has made improvements to its operations accordingly, coverage measurement evaluations and demographic analysis which produce the estimates of differential undercount, show that the barriers persist.

## **References in Appendix A**

Anderson, E. (1989), "Toward the Social Meaning of the Census to the Inner City Poor: Considerations for the Census Undercount." Paper prepared for the U.S. Census Bureau. May 14, 1989.

Aschenbrenner, J. (1975), *Lifelines: Black Families in Chicago*. New York: Holt, Rinehart and Winston.

Aschenbrenner, J. (1991), "A Community-Based Study of the 1990 Census Undercount in a Racially Mixed Area." Ethnographic Evaluation of the 1990 Decennial Census Report #1. Prepared under Joint Statistical Agreement 89-44 with Southern Illinois University at Edwardsville. Bureau of the Census, Washington, D.C.

Bates, N. and Gerber, E. (1998), "Temporary mobility and reporting of usual residence." *Survey Methodology* 24(1):89-98.

Bell, W. (2004), Distant relatives have lower census inclusion probabilities than immediate family. Email message to David Whitford: Chief of Decennial Statistical Studies Division. May 27, 2004.

Bonvillain, N. (1991), "The Census Process at St. Regis Reservation." Ethnographic Exploratory Research Report #3. United States Bureau of the Census, Washington, DC.

Bourgois, P. (1990), "Hypotheses and Ethnography Analysis of Concealment in the Underground Economy: The Economic and Ideological Dynamics of the Census Undercount." Prepared for Joint Statistical Agreement 88-24 in conjunction with Research Institute for the Study of Man New York, NY. United States Bureau of the Census.

Brownrigg, L., and de la Puente, M. (1992, August), "Sociocultural Behaviors Correlated with Census Undercount." Paper presented at Special Session 2 15 at the annual meeting of the American Sociological Association.

Brownrigg, L. and de la Puente, M. (1993), "Alternative Enumeration Methods and Results: Resolution and Resolved Populations by Site". Preliminary Research and Evaluation Memorandum (PREM).

Brownrigg, L. (2003), "Ethnographic Social Network Tracing of Highly Mobile People." Census 2000 Evaluation J.2: Ethnographic Studies, October 16, 2003.

Clark, J. and Moul, D. (2004), "Census 2000 Testing, Experimentation, and Evaluation Program Topic No. 10, Tr-10, Coverage Improvement in the Census 2000 Enumeration," U.S. Census Bureau, Washington DC 20233.

de la Puente, M. (1991), "In Search of the Causes of Differential Census Undercount of Racial and Ethnic Minorities: Overview of Ethnographic Studies of Census Undercount." Paper presented at the 86<sup>th</sup> Annual Meeting of the American Sociological Association, Cincinnati Ohio.

de la Puente, M. (1993a), "Using Ethnography to Explain Why People are Missed or Erroneously Included by the Census: Evidence from Small Ethnographic Studies." Center for Survey Methods Research, U.S. Census Bureau.

de la Puente, M. (1993b), "Why are People Missed or Erroneously Included in the Census: A Summary of Findings from Ethnographic Coverage Reports." In *Proceedings of the 1993 Research Conference on Undercounting Ethnic Populations of the U.S. Census Bureau* (pp. 29-66). Washington, D.C.

de la Puente, M. and Stemper, D. (2003), "The Enumeration of Colonias in Census 2000: Perspectives of Ethnographers and Census Enumerators." *Census 2000 Evaluation J.4: Ethnographic Studies*. September 22, 2003.

Gerber, E. (1990), "A Cognitive Approach to Household Membership Judgments Among Lower Class Blacks," Center for Survey Methods Research, U.S. Bureau of the Census.

Hamid, A. (1992), "Ethnographic Follow-Up of a Predominately African American Population in a Sample Area in Central Harlem, New York City: Behavioral Causes of the Undercount of the 1990 Census." *Ethnographic Evaluation of the 1990 Decennial Census Report #11*. Prepared under Joint Statistical Agreement 89-28 with John Jay College of Criminal Justice in New York City. Bureau of the Census, Washington, D.C.

Hainer, P. (1987), "A Brief and Qualitative Anthropological Study Exploring the Reasons for Census Coverage Error among Low Income Black Households." Report prepared under contract with the Census Bureau. April 8, 1987.

Hudgins, J., Holmes B., and Locke M. (1990), "The Impact of Family Structure among Black Families on the Under Enumeration of Black Males, Parts I and II." Preliminary and Final Reports for the Joint Statistical Agreement 89-6 with Hampton University's Black Family Institute. *Ethnographic Exploratory Research Report 14*. Washington, DC: Bureau of the Census.

Hunter, J., De la Puente M., and Salo, M. (2003), "Comparative Ethnographic Research on Mobile Populations." *Census 2000 Evaluation J.3: Ethnographic Studies*. July 17, 2003.

Lobo, S. (1990), "Oakland's American Indian Community: History, Social Organization, and Factors that Contribute to Census Undercount." Preliminary report for Joint Statistical Agreement 89-19 with Intertribal Friendship House, Oakland, California. *Ethnographic Exploratory Research Report 12*. Washington, D.C: Bureau of the Census.

Martin, E., Brownrigg, L., & Fay, R. (1990), "Results of 1988 Ethnographic Studies of Census Coverage and Plans for 1990." Paper presented to the Census Advisory Committees of the American Statistical Association and Section on Population Statistics at the Joint Advisory Committee Meeting, Alexandria, VA.

Moore, J. (1992), "Determine Extent of Census Undercount Among Certain Rural Creek Indians of Oklahoma." Final Report for Joint Statistical Agreement 89-30 with the University of Oklahoma. *Ethnographic Evaluation of the 1990 Decennial Census Report 10* (March 1992).

Mule, T. (2012), "Census Coverage Measurement Estimation Report: Summary of Estimates of Coverage for Persons in the United States." DSSD 2010 Census Coverage Measurement Memorandum Series #2010-G\_01. May 22, 2012.

Robinson, G., Ahmed B., Das Gupta, P., and Woodrow K. (1993), "Estimation of Population Coverage in the 1990 United States Census Based on Demographic Analysis." *Journal of the American Statistical Association*, Vol. 88, No. 423, pp. 1061-1071.

Robinson, G. and Adlakha A. (2002), "Comparison of A.C.E. Revision II Results with Demographic Analysis." A.C.E. Revision II Estimates Memorandum Series #PP-41, U.S. Census Bureau, December 31, 2002.

Rodriguez, N., and Hagan J. (1991), "Investigating Census Coverage and Content among the Undocumented: Ethnographic Study of Latino Tenants in Houston, Texas." Final Report for Joint Statistical Agreement 89-34 with the University of Houston. Ethnographic Evaluation of the 1990 Decennial Census Report 3 (December 1991).

Schwede, L. (2003), "Complex Households and Relationships in the Decennial Census and in Six Race/Ethnic Groups." Census 2000 Testing and Experimentation Program Ethnographic Study. August 27, 2003. Suitland, MD: U.S. Census Bureau.  
<http://www.census.gov/pred/www/rpts/Complex%20Households%20Final%20Report.pdf>

Schwede, L., Blumberg R., and Chan A. (2006). *Complex Ethnic Households in America*. New York, NY: Rowman & Littlefield Publishers, Inc.

U.S. Census Bureau (1993). "Programs to Improve Coverage in the 1990 Census," 1990 Census of the Population and Housing Evaluation and Research Reports, 1990 CPH-E-3, U.S. Government Printing Office, November 1993.

U.S. Census Bureau (2010), "The Development and Sensitivity Analysis of the 2010 Demographic Analysis Estimates," by the Demographic Analysis Research Team.

U.S. Census Bureau (2012), 2010 Census Coverage Measurement Results News Conference. May 22, 2012. Available at: <http://2010.census.gov/news/press-kits/ccm/ccm.html>.

Valentine, C.B and Valentine B.L. (1971). "Missing Men: A Comparative Methodological Study of Underenumeration and Related Problems." Report prepared under Joint Statistical Agreement with the Brookdale Hospital Center, Washington.


## **Appendix B: Original Study Plan Research Questions**

As mentioned in Section 3, during the preparation of this report, the decision was made that restructuring, combining, and reordering the questions to focus on types and sources of possible coverage error as the foremost theme would be a more effective and parsimonious way to present the findings. The Census Bureau's Decennial Management Division sponsors approved these changes. The original research questions are listed below.

1. How well do the 2020 Census personal visit operations observed—Update Enumerate and Nonresponse Followup—and the Census Coverage Measurement Person Interview work overall and for specific race/ethnic populations?
2. What problems and issues are found with the enumeration methods? With the questionnaires? With the residence rules, rostering, and coverage? What other issues affect coverage?
3. What are the types and sources of possible coverage errors?
4. What are the characteristics of persons and households with possible coverage error? Are coverage errors found more often in complex households (those households with persons other than, or in addition to, one nuclear family of husband, wife, and their joint biological children, one parent with children) than noncomplex households (one person households, stem households with just one parent and just the nuclear family)?
5. What similarities and differences are observed across the race/ethnic sites?
6. What improvements and/or new research can we suggest that might reduce differential miscounts of minority populations in the 2020 Census?

# Appendix C: The Enumerator Questionnaire (EQ) and Information Sheet

OMB No. 0507-0919-C: Approval Expires 12/31/2011



U.S. DEPARTMENT OF COMMERCE  
Economic and Statistics Administration  
U.S. CENSUS BUREAU

## ENUMERATOR QUESTIONNAIRE

Unit ID

|  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|
|  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|

LCO      State      County

|  |  |  |
|--|--|--|
|  |  |  |
|--|--|--|

Tract      Block

|  |  |  |  |
|--|--|--|--|
|  |  |  |  |
|--|--|--|--|

AA      Map Spot

|  |  |  |  |
|--|--|--|--|
|  |  |  |  |
|--|--|--|--|

← APPLY LABEL HERE →

Are there any continuation forms for this address?

Yes → Number of forms

No

---

### RECORD OF CONTACT

| Type   | Mo                   | Day                  | Time  | Outcome  | Type                               | Mo                   | Day                  | Time  | Outcome  |
|--|----------------------|----------------------|---|--|------------------------------------|----------------------|----------------------|---|--|
| <input checked="" type="checkbox"/> Personal | <input type="text"/> | <input type="text"/> | <input type="text"/> : <input type="text"/> | <input type="checkbox"/> a.m.<br><input type="checkbox"/> p.m. | <input type="checkbox"/> Personal  | <input type="text"/> | <input type="text"/> | <input type="text"/> : <input type="text"/> | <input type="checkbox"/> a.m.<br><input type="checkbox"/> p.m. |
| <input type="checkbox"/> Personal            | <input type="text"/> | <input type="text"/> | <input type="text"/> : <input type="text"/> | <input type="checkbox"/> a.m.<br><input type="checkbox"/> p.m. | <input type="checkbox"/> Telephone | <input type="text"/> | <input type="text"/> | <input type="text"/> : <input type="text"/> | <input type="checkbox"/> a.m.<br><input type="checkbox"/> p.m. |
| <input type="checkbox"/> Telephone           | <input type="text"/> | <input type="text"/> | <input type="text"/> : <input type="text"/> | <input type="checkbox"/> a.m.<br><input type="checkbox"/> p.m. | <input type="checkbox"/> Personal  | <input type="text"/> | <input type="text"/> | <input type="text"/> : <input type="text"/> | <input type="checkbox"/> a.m.<br><input type="checkbox"/> p.m. |
| <input type="checkbox"/> Personal            | <input type="text"/> | <input type="text"/> | <input type="text"/> : <input type="text"/> | <input type="checkbox"/> a.m.<br><input type="checkbox"/> p.m. | <input type="checkbox"/> Telephone | <input type="text"/> | <input type="text"/> | <input type="text"/> : <input type="text"/> | <input type="checkbox"/> a.m.<br><input type="checkbox"/> p.m. |
| <input type="checkbox"/> Telephone           | <input type="text"/> | <input type="text"/> | <input type="text"/> : <input type="text"/> | <input type="checkbox"/> a.m.<br><input type="checkbox"/> p.m. | <input type="checkbox"/> Personal  | <input type="text"/> | <input type="text"/> | <input type="text"/> : <input type="text"/> | <input type="checkbox"/> a.m.<br><input type="checkbox"/> p.m. |
|  |                      |                      |   |  | <input type="checkbox"/> Telephone | <input type="text"/> | <input type="text"/> | <input type="text"/> : <input type="text"/> | <input type="checkbox"/> a.m.<br><input type="checkbox"/> p.m. |

OUTCOME CODES: NV = Left Notice of Visit    NC = No Contact    RE = Refusal    CI = Conducted Interview    OT = Other

---

**S1.** Hello, I'm (Name) from the U.S. Census Bureau. (Show ID). Is this (Address)?

Yes - Continue with question S2.

No - Ask: Can you tell me where to find (Address)?  
*END INTERVIEW.*

**S2.** I'm here to complete a Census questionnaire for this address. It should take about 10 minutes. (Hand respondent an Information Sheet.) The first part explains that your answers are confidential. I'll refer to the other parts later. Did you or anyone in this household live or stay here on April 1, 2010?

Yes - Continue with question S3.

No - Skip to question S4.

**S3.** Does someone usually live at this (house/apartment/mobile home), or is this a vacation or seasonal home?

Usually lives here - Skip to question S5.

Vacation or seasonal home or held for occasional use - Skip to "Respondent Information" on back page.

**S4.** (Only ask if no household member lived here on April 1.) On April 1, was this unit vacant, or occupied by a different household?

Vacant - Skip to "Respondent Information" on back page.

Occupied by a different household - Using a knowledgeable respondent, complete this questionnaire for the Census Day household.

Not a housing unit - Skip to "Respondent Information" on back page.

**S5.** We need to count people where they live and sleep most of the time.

Please look at list A. It contains examples of people who should and should not be counted at this place.

Based on these examples, how many people were living or staying in this (house/apartment/mobile home) on April 1?

= Number of people

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



140101

Form D-1(E)  
(12-3-2008)





|   |   |   |
|---|---|---|
| <p>5. Please look at List C. Is (Name) of Hispanic, Latino, or Spanish origin?</p> <p>Read if necessary:<br/>Examples of another Hispanic, Latino, or Spanish origin include Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.</p> <p><input type="checkbox"/> No, not of Hispanic, Latino, or Spanish origin<br/> <input type="checkbox"/> Yes, Mexican, Mexican American, Chicano<br/> <input type="checkbox"/> Yes, Puerto Rican<br/> <input type="checkbox"/> Yes, Cuban<br/> <input type="checkbox"/> Yes, another Hispanic, Latino, or Spanish origin – What is that origin? <input type="text"/></p> | <p>6. Please look at List D and choose one or more races. For this census, Hispanic origin is not a race. What is (Name/s) race?</p> <p>Read if necessary:<br/>Examples of other Asian groups include Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.<br/>Examples of other Pacific Islander groups include Fijian, Tongan, and so on.</p> <p><input type="checkbox"/> White    <input type="checkbox"/> Black, African American, or Negro    <input type="checkbox"/> American Indian or Alaska Native<br/> <input type="checkbox"/> Asian Indian    <input type="checkbox"/> Chinese    <input type="checkbox"/> Filipino<br/> <input type="checkbox"/> Japanese    <input type="checkbox"/> Korean    <input type="checkbox"/> Vietnamese</p> <p><input type="checkbox"/> Native Hawaiian    <input type="checkbox"/> Guamanian or Chamorro    <input type="checkbox"/> Samoan<br/> <input type="checkbox"/> Some other race — What is that group? <input type="text"/></p> <p>What is the name of the enrolled or principal tribe? <input type="text"/></p> <p><input type="checkbox"/> Other Asian — What is that group? <input type="text"/></p> <p><input type="checkbox"/> Other Pacific Islander — What is that group? <input type="text"/></p> | <p>7. Does (Name) sometimes live or stay somewhere else for any of these reasons? – Read response categories. Mark <input checked="" type="checkbox"/> all reasons that apply.</p> <p><input type="checkbox"/> In college housing<br/> <input type="checkbox"/> In the military<br/> <input type="checkbox"/> At a seasonal or second residence<br/> <input type="checkbox"/> For child custody<br/> <input type="checkbox"/> In jail or prison<br/> <input type="checkbox"/> In a nursing home<br/> <input type="checkbox"/> For another reason<br/> <input type="checkbox"/> No</p> |
| <p><input type="checkbox"/> No, not of Hispanic, Latino, or Spanish origin<br/> <input type="checkbox"/> Yes, Mexican, Mexican American, Chicano<br/> <input type="checkbox"/> Yes, Puerto Rican<br/> <input type="checkbox"/> Yes, Cuban<br/> <input type="checkbox"/> Yes, another Hispanic, Latino, or Spanish origin – What is that origin? <input type="text"/></p>  | <p><input type="checkbox"/> White    <input type="checkbox"/> Black, African American, or Negro    <input type="checkbox"/> American Indian or Alaska Native<br/> <input type="checkbox"/> Asian Indian    <input type="checkbox"/> Chinese    <input type="checkbox"/> Filipino<br/> <input type="checkbox"/> Japanese    <input type="checkbox"/> Korean    <input type="checkbox"/> Vietnamese</p> <p><input type="checkbox"/> Native Hawaiian    <input type="checkbox"/> Guamanian or Chamorro    <input type="checkbox"/> Samoan<br/> <input type="checkbox"/> Some other race — What is that group? <input type="text"/></p> <p>What is the name of the enrolled or principal tribe? <input type="text"/></p> <p><input type="checkbox"/> Other Asian — What is that group? <input type="text"/></p> <p><input type="checkbox"/> Other Pacific Islander — What is that group? <input type="text"/></p>   | <p><input type="checkbox"/> In college housing<br/> <input type="checkbox"/> In the military<br/> <input type="checkbox"/> At a seasonal or second residence<br/> <input type="checkbox"/> For child custody<br/> <input type="checkbox"/> In jail or prison<br/> <input type="checkbox"/> In a nursing home<br/> <input type="checkbox"/> For another reason<br/> <input type="checkbox"/> No</p>  |
| <p><input type="checkbox"/> No, not of Hispanic, Latino, or Spanish origin<br/> <input type="checkbox"/> Yes, Mexican, Mexican American, Chicano<br/> <input type="checkbox"/> Yes, Puerto Rican<br/> <input type="checkbox"/> Yes, Cuban<br/> <input type="checkbox"/> Yes, another Hispanic, Latino, or Spanish origin – What is that origin? <input type="text"/></p>  | <p><input type="checkbox"/> White    <input type="checkbox"/> Black, African American, or Negro    <input type="checkbox"/> American Indian or Alaska Native<br/> <input type="checkbox"/> Asian Indian    <input type="checkbox"/> Chinese    <input type="checkbox"/> Filipino<br/> <input type="checkbox"/> Japanese    <input type="checkbox"/> Korean    <input type="checkbox"/> Vietnamese</p> <p><input type="checkbox"/> Native Hawaiian    <input type="checkbox"/> Guamanian or Chamorro    <input type="checkbox"/> Samoan<br/> <input type="checkbox"/> Some other race — What is that group? <input type="text"/></p> <p>What is the name of the enrolled or principal tribe? <input type="text"/></p> <p><input type="checkbox"/> Other Asian — What is that group? <input type="text"/></p> <p><input type="checkbox"/> Other Pacific Islander — What is that group? <input type="text"/></p>   | <p><input type="checkbox"/> In college housing<br/> <input type="checkbox"/> In the military<br/> <input type="checkbox"/> At a seasonal or second residence<br/> <input type="checkbox"/> For child custody<br/> <input type="checkbox"/> In jail or prison<br/> <input type="checkbox"/> In a nursing home<br/> <input type="checkbox"/> For another reason<br/> <input type="checkbox"/> No</p>  |
| <p><input type="checkbox"/> No, not of Hispanic, Latino, or Spanish origin<br/> <input type="checkbox"/> Yes, Mexican, Mexican American, Chicano<br/> <input type="checkbox"/> Yes, Puerto Rican<br/> <input type="checkbox"/> Yes, Cuban<br/> <input type="checkbox"/> Yes, another Hispanic, Latino, or Spanish origin – What is that origin? <input type="text"/></p>  | <p><input type="checkbox"/> White    <input type="checkbox"/> Black, African American, or Negro    <input type="checkbox"/> American Indian or Alaska Native<br/> <input type="checkbox"/> Asian Indian    <input type="checkbox"/> Chinese    <input type="checkbox"/> Filipino<br/> <input type="checkbox"/> Japanese    <input type="checkbox"/> Korean    <input type="checkbox"/> Vietnamese</p> <p><input type="checkbox"/> Native Hawaiian    <input type="checkbox"/> Guamanian or Chamorro    <input type="checkbox"/> Samoan<br/> <input type="checkbox"/> Some other race — What is that group? <input type="text"/></p> <p>What is the name of the enrolled or principal tribe? <input type="text"/></p> <p><input type="checkbox"/> Other Asian — What is that group? <input type="text"/></p> <p><input type="checkbox"/> Other Pacific Islander — What is that group? <input type="text"/></p>   | <p><input type="checkbox"/> In college housing<br/> <input type="checkbox"/> In the military<br/> <input type="checkbox"/> At a seasonal or second residence<br/> <input type="checkbox"/> For child custody<br/> <input type="checkbox"/> In jail or prison<br/> <input type="checkbox"/> In a nursing home<br/> <input type="checkbox"/> For another reason<br/> <input type="checkbox"/> No</p>  |
| <p><input type="checkbox"/> No, not of Hispanic, Latino, or Spanish origin<br/> <input type="checkbox"/> Yes, Mexican, Mexican American, Chicano<br/> <input type="checkbox"/> Yes, Puerto Rican<br/> <input type="checkbox"/> Yes, Cuban<br/> <input type="checkbox"/> Yes, another Hispanic, Latino, or Spanish origin – What is that origin? <input type="text"/></p>  | <p><input type="checkbox"/> White    <input type="checkbox"/> Black, African American, or Negro    <input type="checkbox"/> American Indian or Alaska Native<br/> <input type="checkbox"/> Asian Indian    <input type="checkbox"/> Chinese    <input type="checkbox"/> Filipino<br/> <input type="checkbox"/> Japanese    <input type="checkbox"/> Korean    <input type="checkbox"/> Vietnamese</p> <p><input type="checkbox"/> Native Hawaiian    <input type="checkbox"/> Guamanian or Chamorro    <input type="checkbox"/> Samoan<br/> <input type="checkbox"/> Some other race — What is that group? <input type="text"/></p> <p>What is the name of the enrolled or principal tribe? <input type="text"/></p> <p><input type="checkbox"/> Other Asian — What is that group? <input type="text"/></p> <p><input type="checkbox"/> Other Pacific Islander — What is that group? <input type="text"/></p>   | <p><input type="checkbox"/> In college housing<br/> <input type="checkbox"/> In the military<br/> <input type="checkbox"/> At a seasonal or second residence<br/> <input type="checkbox"/> For child custody<br/> <input type="checkbox"/> In jail or prison<br/> <input type="checkbox"/> In a nursing home<br/> <input type="checkbox"/> For another reason<br/> <input type="checkbox"/> No</p>  |

H1. We do not want to miss any people who might have been staying here on April 1.  
 Were there any additional people that you didn't mention, for example:

- Babys?  Yes  No
- Foster children?  Yes  No
- Any other relatives?  Yes  No
- Roommates?  Yes  No
- Any other nonrelatives?  Yes  No
- How about anyone else staying here on April 1 who had no permanent place to live?  Yes  No

If yes to any category, ask: What is that person's name?

|                      |                      |
|----------------------|----------------------|
| First Name           | Last Name            |
| <input type="text"/> | <input type="text"/> |

Anyone else?

|                      |                      |
|----------------------|----------------------|
| First Name           | Last Name            |
| <input type="text"/> | <input type="text"/> |

Do not list any people recorded for this question on the inside pages or on a continuation form.

H2. Do you or does someone in this household own this (house/apartment/mobile home) with a mortgage or loan, including home equity loans, own it free and clear, rent it, or occupy it without having to pay rent?

- Own with a mortgage or loan (including home equity loans)
- Own free and clear (without a mortgage or loan)
- Rent
- Occupy without payment of rent

H3. If there is not an address listed above or if the address listed includes only a location description, ask: What is the address of this unit?

|                      |                                    |                      |
|----------------------|------------------------------------|----------------------|
| House number         | Street name or rural route address |                      |
| <input type="text"/> | <input type="text"/>               |                      |
| Apartment number     | <input type="text"/>               |                      |
| <input type="text"/> |                                    |                      |
| City                 | State                              | ZIP Code             |
| <input type="text"/> | <input type="text"/>               | <input type="text"/> |

Go to Respondent Information on back page.

**NOTES**

|  |  |  |  |  |
|--|--|--|--|--|
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**RESPONDENT INFORMATION**

**R1. (Just or empty) What is your name?**

First Name  M

Last Name

Address of proxy

**R2. What is your phone number and best time to call?**

Area Code  Number  -

Day  Evening  Other

**R3. Respondent Type -**

Household member - Lived here on April 1, 2010

Household member - Moved in after April 1, 2010

Neighbor or other proxy

**INTERVIEW SUMMARY**

**A. Unit Status on April 1, 2010**

Occupied

Vacant - regular

Vacant - usual home elsewhere

Demolished/burned/outboarded locale

Nonresidential

Empty mobile home/trailer site

Uninhabitable (open to elements, condemned, under construction)

Duplicate - record ID of Dup is:

**B. If vacant unit: Which category best describes this vacant unit as of April 1, 2010? (Select all that apply.)**

For rent

Rented, not occupied

For sale only

Sold, not occupied

For seasonal, recreational or occasional use

For migrant workers

Other vacant

**C. Number of people listed on form(s) =**

or - or - Total people

or - Vacant

or - Delete

or - PCP unknown

**D. What language was the majority of the interview conducted in?**

English

Spanish

Other - Specify language number from list out -

**E.**  USE  MOV  FI  REP

**F.**  CD  RSP  YOC

**G.**  L  J  K

**H.**  JIC1  JIC2

**I.**  L  M

**CERTIFICATION**

I certify that the entries I have made on this questionnaire are true and correct to the best of my knowledge.

Enumerator's signature  Employee ID

Month  Day

Crew Leader's initials

CLD number

Month  Day

# The 2010 Census NRFU Information Sheet



## List A

**Your Answers Are Confidential**

Your answers are confidential and protected by law. All U.S. Census Bureau employees have taken an oath and are subject to a jail term, a fine, or both if they disclose ANY information that could identify you or your household. Your answers will only be used for statistical purposes, and no other purpose. As allowed by law, your census data becomes public after 72 years. This information can be used for family history and other types of historical research.

You are required by law to provide the information requested. These federal laws are found in the United States Code, Title 13 (Sections 9, 141, 193, 214, and 221) and Title 44 (Section 2108). Please visit our Web site at <[www.census.gov/2010census](http://www.census.gov/2010census)> and click on "Protecting Your Answers" to learn more about our privacy policy and data protection.

Thank you for your cooperation. The U.S. Census Bureau appreciates your help.

**WHO TO COUNT ON APRIL 1<sup>st</sup>**

**We need to count people where they live and sleep most of the time.**

| <u>Do NOT include:</u>   | <u>Do include:</u>  |
|--|---|
| <ul style="list-style-type: none"> <li>College students who live away from this address most of the year</li> <li>Armed Forces personnel who live away</li> <li>People in a nursing home, mental hospital, etc. on April 1, 2010</li> <li>People in jail, prison, detention facility, etc. on April 1, 2010</li> </ul> | <ul style="list-style-type: none"> <li>Babies and children living here, including foster children</li> <li>Roommates</li> <li>Boarders</li> <li>People staying here on April 1, 2010 who have no permanent place to live</li> </ul> |

If you have any comments concerning the time it takes to complete this form or any other aspect of the collection, send it to: Paperwork Reduction Project 0607-0919-C, U.S. Census Bureau, AMSD-3K138, 4600 Silver Hill Road, Washington, DC 20233. You may e-mail comments to <[Paperwork@census.gov](mailto:Paperwork@census.gov)>; use "Paperwork Project 0607-0919-C" as the subject.

Respondents are not required to respond to any information collection unless a valid approval number has been assigned by the Office of Management and Budget. The approval number for the 2010 Census is: OMB No. 0607-0919-C; Approval Expires 12/31/2011.

D-1(F) (9-20-2009)

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

## List B

**RELATIONSHIP**

Husband or wife

Biological son or daughter

Adopted son or daughter

Stepson or stepdaughter

Brother or sister

Father or mother

Grandchild

Parent-in-law

Son-in-law or daughter-in-law

Other relative

Roomer or boarder

Housemate or roommate

Unmarried partner

Other nonrelative

D-1(F) (9-20-2009)

## List C

**HISPANIC, LATINO, OR SPANISH ORIGIN**

No, not of Hispanic, Latino, or Spanish origin

Yes, Mexican, Mexican American, or Chicano

Yes, Puerto Rican

Yes, Cuban

Yes, another Hispanic, Latino, or Spanish origin – *For example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.*

## List D

**RACE**  
*(Choose one or more races.)*

White

Black, African American, or Negro

American Indian or Alaska Native

Asian Indian

Chinese

Filipino

Japanese

Korean

Vietnamese

Other Asian – *For example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.*

Native Hawaiian

Guamanian or Chamorro

Samoan

Other Pacific Islander – *For example, Fijian, Tongan, and so on.*

Some other race