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

June 29, 2012

2010 CENSUS PLANNING MEMORANDA SERIES

No. 208

MEMORANDUM FOR The Distribution List

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

Subject: 2010 Census Coverage Measurement Person Interview Operation Assessment

Attached is the 2010 Census Coverage Measurement Person Interview Operation Assessment. The Quality Process for the 2010 Census Test Evaluations, Experiments, and Assessments was applied to the methodology development and review process. The report is sound and appropriate for completeness and accuracy.

If you have any questions about this document, please contact Kyra Linse at (301) 763-9280.

Attachment

April 2, 2012

2010 Census Coverage Measurement Person Interview Operation Assessment

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

Final Report

Kyra Linse, Ashley C. Argarin

Decennial Statistical Studies Division,
Decennial Management Division

TABLE OF CONTENTS

List of Tables

List of Figures

Executive Summary

In this assessment of the Census Coverage Measurement Person Interview Operations, we document the operational implementation and present results to the assessment research questions to ascertain the operations' effectiveness and identify any areas that may need improvement. The goal of Person Interview was to list the household roster of the sample address and collect information to determine where each person should have been counted on Census Day (April 1, 2010). The final data from the Person Interview Operation are inputs into the matching operation of the census persons enumerated in the Census Coverage Measurement sample areas and in the final estimation of coverage for the 2010 Census.

The Census Coverage Measurement Person Interview Operation was conducted from August 14, 2010 through October 4, 2010. The purpose of the Person Interview was to obtain a roster and information about the residents of the sample-housing unit at the time of the interview. This includes nonmovers and people who may have moved into the selected housing unit since Census Day, known as inmovers. In addition, the Person Interview collected information about persons who moved out of the sample-housing unit between Census Day and the time of the Census Coverage Measurement interview (outmovers). The demographic information collected for each person included name, sex, age, date of birth, race, relationship, and Hispanic origin. It also collected information to determine where each sample unit's current resident was living on Census Day and the new address for each outmover. The Person Interview also collected information to determine if there were any other alternate addresses where any of the people listed may have been counted in the census and information necessary to geocode the alternate addresses for future operations.. The Census Coverage Measurement Person Interview was conducted by personal visit interviewers using a computer-assisted data collection instrument on a laptop computer. A telephone wording path was also provided to support instances when respondents requested an interview via telephone. In addition, there was a Spanish translation available.

The other component to the operation was a quality check of the operation, known as the Census Coverage Measurement Person Interview Reinterview operation. This operation was conducted to check the quality of the original Person Interview work done with the intent to detect and attempt to correct any instances of falsification. This involved conducting an independent Person Interview Reinterview with a sample of households to verify that the interviewer actually conducted the Person Interview and followed appropriate procedures. In situations where the Person Interview Reinterview instrument determined from the responses that a Person Interview might not have been conducted, a full Person Interview was conducted using the Person Interview Reinterview instrument. After a reinterviewer completed a case, the data from the two interviews were compared through computer and clerical matching, as well as a Regional Census Center review, to determine if the original interviewer conducted the interview correctly, followed proper procedures, and collected complete data. If it was determined that the case was falsified or incorrect, then the data collected in Person Interview Reinterview were used as the final version of the interview instead of the original Person Interview.

We list below the main conclusions from the Person Interview assessment:

- Person Interview had an overall response rate of 98.4 percent.

- The 2010 Census Coverage Measurement Person Interview operation was under budget by $8,333,640 (25.98 percent).
- The missing data rate was under 2.78 percent for each of the residence questions.
- We knew if the person should be counted at the Census Day residence for over 98 percent of people for whom we collected data.
- We collected 60,950 alternate addresses in Person Interview and only 7.74 percent of the addresses did not provide enough data to establish a location.

In addition, this report answers the following questions:

How was the workload distributed over the length of the survey? - The 2010 Person Interview was conducted in 186,766 selected housing units. The workload was delivered in one delivery to all of the 12 Regional Census Centers in time to begin on August 14, 2010. The operation was very successful in completing all work on time. Within the four week mark, by September 10, 2010, 91.00 percent of cases had been completed and the last two weeks were used to complete the hard to contact cases.

How well did Census Coverage Measurement Person Interview conform to the expected schedule and cost? – Person Interview and Person Interview Reinterview were both finished on the baseline end dates.

Overall, the Person Interview operations were under budget by $8,333,640 (25.98 percent). Production Person Interview was under budget by $7,050,080 (27.45 percent) and Person Interview Reinterview was under budget by $1,283,561 (20.08 percent). The costs for field work, training, and Per Diem were under budget; however mileage costs were over budget. Still, lower than budgeted costs in all other areas offset these extra mileage costs.

How did Field staffing and training plans meet the needs for Person Interview production? Authorized staffing levels were more than sufficient to perform and complete both Person Interview production and Person Interview Reinterview; Person Interview hired only 85.28 percent of their authorized staff and Reinterview hired only 93.79 percent of their authorized staff. In addition, there was little sign of attrition with only 9.54 percent (460) of Person Interview staff being trained after the operation started and only 3.31 percent (32) of Reinterview staff being trained after the Reinterview operation started.

What was the response rate and cooperation rate for Person Interview? – Overall, the final response rate for production was 98.43 percent and the final cooperation rate was 98.68 percent. The response rate for each Regional Census Center was 97.04 percent or higher except for the New York Regional Census Center at 95.44 percent. All Regional Census Centers followed a similar pattern for their cooperation rates.

How effective was the automation of the Person Interview? – Automating the Person Interview and Person Interview Reinterview instruments required extensive testing within a very short time but led to a very high quality operation and data.

Overall, automation led to better data by eliminating potential interviewer error in navigating through the interview and later potential keying or scanning errors if paper instruments had been used.

The Coverage Measurement Operations Control System allowed not only very detailed tracking of cases in the field, but it also allowed almost real-time daily review of data at Headquarters to give feedback to Field staff. It also allowed us to identify and correct some instrument or procedural problems, immediately eliminating or reducing poor quality cases. Overall, only 23 problems were found and no single problem affected more than 100 cases.

How many of the production cases failed the quality check and needed to be replaced with the Reinterview cases? – Approximately 16.50 percent (30,745) of cases from Person Interview were sampled for Person Interview Reinterview. Out of those, there were 812 cases that did not pass the quality check review, meaning they were marked either incorrect or possibly incorrect. For 789 of those, the Reinterview case was selected to replace the Person Interview. Of those replaced, 669 (84.79 percent) were selected to replace the original interview because the information in the interview was confirmed wrong, but it was determined that it was due to a mistake by the interviewer and not intentional falsification. Overall, 107 cases were confirmed to be "Hard Fails" (i.e. confirmed as intentionally falsified). These cases came from 19 different interviewers.

How effective was Person Interview at rostering people? –Person Interview was designed to list all people who lived or stayed at the sample address, including people tenuously attached to the address. To encourage the respondent to list anyone that stayed at the address, in addition to the original roster, we also had four probes that asked about populations that the respondent might not initially consider as part of their household. Once we had that roster, we also collected any individual outmovers who have left the sample address since around Census Day. If the entire household had left (e.g., the house is currently vacant, not a housing unit, or has a whole household that was not staying at the sample address on Census Day), then we also collected a separate roster of those people who have moved out. This allowed us to match to any possible person who could have been listed in the Census at this address.

As expected most people (93.65 percent) were added in the main roster, but a total of 2.47 percent of the people were added through the four probes and 3.76 percent of people listed were collected by the two outmover rosters for a total of 423,242 people. The average household size for complete interviews was 2.69 people.

The data collected for each person were very complete. Overall, 98.72 percent (417,840) of the people had some part of the name reported and could be used in matching. For collecting demographics, there was a very low missing data rate with age reporting having the highest missing rate at 6.71 percent and all others under 2.40 percent.

Probably the most interesting finding was that when comparing related households to nonrelated households (ignoring single person households), the nonrelated households did average slightly more people per household than related households; 3.32 and 3.22 respectively.

How effective was Person Interview at collecting alternate addresses? -- In order to determine if a person is a census erroneous enumeration or a duplicate, we need to collect alternate addresses where people could have been counted on Census Day other than the sample address. Similar to collecting a roster, the interview asks about different types of alternate addresses to make sure we have all the possible addresses where each person could have been counted, such as a college address, a relative address in shared custody situations, or second homes.

Person Interview appears to have done a very good job at collecting alternate addresses. Overall, 60,590 alternate addresses were collected. For all cases, 23.15 percent (43,253) had at least one alternate address and 18.80 percent (79,590) of all people collected reported an alternate address. For cases that had an alternate address reported, the average number of alternate addresses within a housing unit was 1.40 addresses with the maximum number of unique addresses collected in a single case being 18. For persons with an alternate address, the average was 1.10 addresses per person with the maximum of five unique alternate addresses per person. The average number of people connected to an address is 1.45 with one address reported actually being connected to 30 people.

Of the Stateside alternate addresses collected, 46.34 percent were complete. Another 15.49 percent, while not complete, contained enough address components that they were likely to be geocoded. In addition to the address, we collected three other components about the alternate addresses; nearby landmarks, cross streets, and neighbors. These questions, while highly useful during geocoding, ended up being very sensitive to the respondent. Landmarks and cross streets were reported only around 46 percent of the time and neighbors were reported only 13.88 percent of the time when alternate addresses were collected.

How did Person Interview do at collecting the timing of when people were living or staying at the Sample Address or Alternate Address? – In addition to collecting an alternate address, we needed to collect the timing about when that person was at each address. This is a complicated task with all the different types of living situations and the different residence rules based on the type of unit. We collected move date for movers. We collected where a person was most of the time for cyclers (people who go between two or more addresses regularly). If a person had a Group Quarters (GQ) address, we also collected where the person was on Census Day. For each of these categories, the instances of being unable to report the timing of where they were staying was very small. A respondent could not report if the move was before or after April 1, 2010 only 1.81 percent of the time. A respondent could not report where a cycler spent most of the time only 3.25 percent of the time. For those who reported staying in a GQ at some point in 2010, we did have a very high missing rate of 50.76 percent, if a person was in a GQ on Census Day, but this was mainly caused by a flaw in the instrument and not due to nonresponse. Overall, though, the instrument was very successful with only 1.17 percent of people missing the timing needed to determine census residence.

How effective was Person Interview at collecting information to determine where people should be counted on Census Day? – Between the address reported and timing collected, we want to be able to determine where a person should be counted on Census Day. Person Interview assigned an interim residence code at the end of the survey period through the post processing system of collected data. This code was a starting point in Person Matching. Post

Processing assigned a good residence status code (i.e., not Review or Unresolved) to most of the persons reported in Person Interview (80.62 percent), while 18.43 percent of people were sent to clerical review to assign a residence status code. While this percent is high, it was expected because we were conservative in assigning the Review code to most people in complex living situations, so their case would be clerically reviewed.

What were the missing data rates for residence questions? What other questions had a high missing data rate? – To be sure we have accurately collected all possible alternate addresses; there are certain questions within the Person Interview that must be answered to assign a residence code with confidence. These include the main residence questions and the questions determining the timing of staying at the addresses. We reviewed these questions for their missing data rates, which include "Don't Know" and "Refused" answers and exiting the interview before finishing. We also reviewed the rates by the type of respondent, sample occupant compared to proxy.

The missing data rates vary between 0.46 percent and 2.77 percent overall for all the main questions except one (Census Day Status for Whole Household of Outmovers). As expected, the proxy respondents were not able to provide as much detail as the current occupants, but the extremely high missing rate of over 30 percent observed on many of the key questions when proxies are responding was surprising and concerning. On the other hand, it is very good to see that for most key questions the missing rate is less than one percent when talking to a household occupant.

What impact does using proxies have on data quality? – To get a higher response rate, Person Interview allowed proxy respondents. The only time a proxy was not allowed was if the sample address occupant refused. In the end, 19.37 percent (36,175) of the total cases were completed by a proxy respondent; however, 80.16 percent (28,999) of those were for "Vacant" and "Not a Housing Unit" situations where a sample occupant was not available. Therefore, proxies were used only 4.62 percent (7,176) of the time for complete interviews, i.e. excluding the vacant and not a housing unit cases. While reviewing the data from proxies, we were concerned with the completeness of the data they provided.

- The missing rate for the key residence questions is over 27 percent on many of the questions. The missing rate for demographic items ranges from 5.53 percent for sex to 52.20 percent for age.
- The average household size reported by proxies is 1.76 people compared to 2.73 people for sample occupant respondents.
- About 43.03 percent of the alternate addresses reported by proxies did not contain any component of an address compared to only 4.15 percent for occupant respondents.

Interestingly though, 19.78 percent of people listed by proxies for a complete interview had an alternate address compared to 17.80 percent for sample address occupants.

What impact did the Nonsampling Error Reduction Initiatives have on the Person Interview operation? – Overall, the nonsampling error reduction initiatives are very difficult to analyze for specific quality results, but the common understanding of most stakeholders is that they helped ensure data quality and we did not see any negative impacts from the changes. The

most valuable initiative was most likely extending Person Reinterview for one week. In the week added, 385 more Reinterview cases were selected to be reviewed. Most (135) were selected because an interviewer was found to have been falsifying data. If this extra week had not been added, these cases would not have been reinterviewed.

The main lessons learned and recommendations from the Person Interview operation were:

- Many respondents prefer to finish an interview by telephone after initial contact. Both the regional staff and interviewers conducted phone interviews when the respondent called back from a Notice of Visit left at the address. This is what most respondents requested when they contacted either the Regional Census Centers or the interviewer.. We should look into adding a telephone phase or some sort of self-response option for the future.
- Training not only needs hands-on practice interviews for learning the instrument content and automation, but also needs realistic difficult interaction scenarios to better prepare our interviewers for gaining cooperation at the door. If our sample continues to be the same size as 2010, we should consider using ongoing current demographic survey interviewers from other Census Bureau operations who already have this ability and will not need as much training.
- Guidelines on when an interviewer should use a proxy as a contact were difficult to follow not only for the interviewer, but also for some of the regional staff. Use of proxies in the future needs to be reviewed looking at the quality of the data compared with the cost/difficulties. If we do allow proxy respondents again, the guidelines need to be more consistent and have fewer exceptions, or be completely automated as part of the instrument.
- Serious respondent fatigue due to other Census and Census Coverage Measurement operations was occurring by the time Person Interview occurred. We need to review to see if there are ways to limit the number of contacts for the same people resulting from all of the Census and Census Coverage Measurement operations.
- Census Coverage Measurement needs to revisit the timing for Person Interview, that is move it as close to Census Day as possible. This should help reduce respondent reluctance to participate (they will not think Census is over) and should lead to higher data quality by reducing the number of people who will have a change in their living situation since Census Day.

1. Introduction

1.1 Scope

The primary purpose of this assessment is to provide data and assess what happened during the 2010 Census Coverage Measurement (CCM) Person Interview (PI) Operation. This assessment will produce valuable data for the next planning cycle for the 2020 Census and provide information on the success of the operation and impacts to the 2010 CCM Program.

The CCM PI Operational Assessment will document workloads and processing as well as lessons learned for all aspects of the PI Operation including fieldwork, Cost and Progress Reporting, Person Interview Data Output, PI Reinterview (RI) Data Outputs, and any debriefing reports.

1.2 Intended Audience

This document is intended to be a review of the 2010 operation and should be used by anyone wanting to know about the 2010 operation's success and efficiency. It should also be used by anyone working on the operational development of 2020 CCM for guidance on issues and successes.

2. Background

2.1 Overview of Census Coverage Measurement and 2010 Census Coverage Measurement Person Interview

The purpose of the 2010 CCM Program is to evaluate coverage error in the 2010 Census. The CCM is designed to measure the coverage of housing units and persons in the United States, including Puerto Rico. However, for 2010, Remote Alaska, GQ and persons residing in GQs are excluded from the CCM program sample. The CCM provides estimates of the net coverage error and the components of census coverage, including omissions and erroneous inclusions[1]. Since the CCM is an evaluation, the results do not affect the 2010 Census.

The 2010 CCM is a large, complex program conducted independently of the census. The CCM includes five sampling activities, five data collection activities, three computer and three clerical matching activities and the estimation component. There are seven separate operational groups that cover the entire CCM process:

- CCM Sample Design Operation
- CCM Independent Listing Operation
- CCM Initial Housing Unit Matching and Followup Operational Group
- CCM Person Interview Operation
- CCM Person Matching and Followup Operational Group
- CCM Final Housing Unit Matching and Followup Operational Group, and

[1] In terms of people, omissions are people who were not included in the Census at the correct location. Erroneous Enumerations are people who were listed by the Census at the location more than once or should not have been included at all.

- CCM Estimation Operation

There is a separate assessment of each operational group. See Section 6 for each of the official assessment names. This paper will focus only on the assessment of the Person Interview Operation.

For each CCM sample block cluster remaining in sample after the CCM sample reduction (refer to Section 2.3), we conducted a CCM PI for the independent P Sample units and a sample of census housing units[2]. For purposes of this assessment, we refer to all housing units sent to PI as PI sample housing units. The CCM PI was conducted by personal visit using a Computer-Assisted Personal Interview (CAPI) instrument on a laptop computer. A telephone wording path was also provided to support instances when respondents requested an interview via telephone. CCM PI was conducted from August 4 to October 2, 2010. The purpose of the PI was to obtain information about the residents of the sample housing unit at the time of the interview. This includes nonmovers and people who may have moved into the selected housing unit since Census Day (April 1, 2010), known as inmovers. In addition, PI collected information about people who were living at the sample unit around Census Day to identify persons who moved out of the sample housing unit between Census Day and the time of the CCM interview (outmovers). The demographic information collected for each person includes name, sex, age, date of birth, race, relationship, and Hispanic origin. The PI also collected information to determine if there were any other alternate addresses where any of the people listed may have been counted in the Census and information necessary to geocode the alternate addresses for future operations. It also collected information to determine where each current resident was living on Census Day and the new address where each outmover currently lives.

The other component to the operation was a quality check known as the CCM PI RI operation. Its purpose was to check the quality of the work done in PI and detect any instances of falsification. This involved conducting an independent PI RI with a sample of households to verify that the interviewer actually conducted the PI interview and followed appropriate procedures. In situations where the PI RI instrument determined from the responses that a PI may not have been conducted, a full PI interview was conducted using the PI RI instrument. After a reinterviewer completed a case, PI and PI RI data were compared through computer and clerical matching, as well as a Regional Census Center (RCC) review, to determine if the original interviewer conducted the interview correctly, followed proper procedures, and collected complete data. If PI RI determined that the case was falsified or incorrect, then the data collected in PI RI were used as the final version of the interview instead of the original PI.

During the operation, the RCC sent all versions of an interview deemed to be accurate to Headquarters. The Processing Systems Development Branch (PSDB) in the Decennial Statistical Studies Division (DSSD) received and performed post-processing of all collected data. The two main purposes were to: 1) select the final version for each case, and 2) select the final persons rostered and assign a preliminary computer-generated residence status code[3] to each person at

[2] The PI interview workload includes interviews at census-only housing units in addition to CCM sample housing units. Census-only housing units are either addresses in the sample block cluster that are valid housing units missed by CCM or census addresses geocoded to the block cluster that CCM believe are located outside the block cluster.

[3] Every person listed in PI is assigned a residence status code. This code indicates the type of living situation for each person based on the answers given in PI, such that each person is classified as either a nonmover, inmover or

the sample unit. This review also included a daily component of monitoring the operation for any technical or procedural issues.

2.2 History of the Coverage Measurement Person Interview Operation

2.2.1 Census 2000

In 2000, the CCM operation was known as the Accuracy and Coverage Evaluation (A.C.E). The major objective of the A.C.E PI operation was to obtain interviews at housing units in the sample block clusters. There were two phases to the A.C.E. PI operation; telephone and personal visit. Sample addresses that provided a telephone number on their Census questionnaire and were not small multi-unit housing units or not units at rural addresses were interviewed during the telephone phase. Sample addresses that did not meet the criteria or addresses for which a phone interview was not obtained by the end of the telephone phase were interviewed during the personal visit phase. Both phases were conducted using CAPI PI instruments gathering information to determine the person's residence status at the time of the A.C.E. interview and the Census Day status for the people at the sample address. The operation began with the telephone phase on May 8, 2000, and ended with the personal visit phase on September 1, 2000 (Whitford 2001 and Stuart 2003).

2.2.2 2004 Development and 2006 Census Test

Starting in 2004, DSSD along with Statistical Research Division (SRD) started testing and developing content for the PI instrument based on the new goal to collect the data necessary to measure the components of census coverage. The development of the PI updates focused on the following things:

- Creating a better way to build a roster and the best probe to obtain information for all people attached to the household.
- Collecting alternate addresses and better identify the type of address collected.
- Collecting as complete an address as possible along with collecting additional information about the addresses to be able to locate them during geocoding and future operations.
- Collecting demographic information in the most efficient way and as consistently with Census operations as possible.

The PI team used many techniques to determine these methods: research and discussion, usability testing, and cognitive testing. These will not be described in detail here. For most information about this development, please see Section 9 for a list of references.

A major goal of CCM in the 2006 Census Test was to test new methodology for improving the determination of Census Day residence. New data collection methods were designed for the 2006 CCM PI to allow us to determine whether census enumerations were counted correctly[4]

---

outmover. A temporary code is assigned at the end of PI and a Final one is assigned at the end of CCM estimation. For more on Residence Status Codes, see Section 5.12.

[4] For 2006, an enumeration was considered "correct" for net error estimation if the person should have been counted in the block or surrounding blocks on Census Day. An enumeration was considered "correct" for components of

with sufficient accuracy to support the estimation of the components of census coverage at various levels of geography; including erroneous enumerations and omissions. The data collection effort was designed to use information on persons who live at the sample address on Interview Day and certain outmovers, known as PES (Post-Enumeration Survey)-B+ methodology[5]. This was similar to the methodology used in the 1990 PES, but differed from the method used in the 2000 A.C.E. The results from the 2006 test and behavior coding from that test were used to better refine the instrument for the 2009 test. (Whitford 2007 and Adams and Nichols 2007)

2.2.3 2009 Operational Test

To reduce risk to the 2010 Census, CCM was removed from the Field Data Collection Automation (FDCA) contract and the development of the automated interviewing instruments for both the 2008 CCM test and the 2010 PI and PI RI was assigned to the Census Bureau's Technologies Management Office (TMO). As a result of this reassignment, the platform for the PI and PI RI instruments switched from a hand-held computer to a laptop. TMO was also assigned with developing the systems and software needed to control and manage the PI, PI RI, and all other 2010 CCM data collection activities, via the Coverage Measurement Operations Control System (CMOCS). The PI test was conducted from April 17 to May 22, 2009, and included both the automated PI instrument and an automated PI RI instrument on laptops. The interviews were modeled off the instrument developed in 2006 with approved changes based on results of the test. The data collected in the two dress rehearsal sites; Stockton, CA and Fayetteville, NC; were used for purposes of software testing and systems development for 2010 CCM applications. No formal assessments were completed before 2010 production due to timing constraints, but data were reviewed and lessons learned were collected to implement necessary changes in time for the 2010 production. Due to the tight timing between the 2009 test and production, it was agreed only major necessary changes would be implemented for 2010. The full review of the 2009 PI was completed around the time of 2010 production (Stone et al., 2010).

2.3 Nonsampling Error Reduction Initiatives for Census Coverage Measurement for 2010

In September 2009, the Census Bureau announced an initiative to reduce nonsampling error in the CCM program. In order to make the changes cost neutral, CCM managers reduced the sample size for operations *after* the CCM Independent Listing, which was almost complete nationwide, and diverted the resulting funds towards approaches to reduce the nonsampling error. The following initiatives were implemented for PI:

- **Higher field work reinterview rates** – By increasing our reinterview rates for our field operations, we hoped to ensure a higher quality product.

---

census coverage if the person should have been counted in the census and was counted once and only once within the specified type of geographic area, which could be nation, state, or county.

[5] People who have moved out of the sample address to somewhere that is not eligible for CCM are included in PES-B+. This includes people who have moved outside of the country, moved to group quarters, have died, or are experiencing homelessness.

- **Adding training modules to interviewer training** – Several modules were added to interviewer training, including more localized training scenarios, training on situations due to current economic conditions (squatters, temporary movers, etc.), and training on probing for other residences.
- **Smaller employee-to-supervisor ratios** – We redefined staffing parameters so fewer lower-level field staff were assigned to each supervisor. We hoped to ensure a greater control over the quality of the field work by allowing more monitoring of work at each level.
- **Extending PI RI for one week** – Due to scheduling constraints, the original plan for reinterviewing was to stop sampling one week before PI production finished to allow time to complete the reinterview cases without delaying final data. With a smaller workload flowing to the later operations, we were able to extend the time for PI RI to allow sampling to continue throughout the duration of PI. (Whitford, 2009)
- **Extra Observation of Interviewers** – To make sure interviewers kept quality in the forefront, we decided to do an extra observation in addition to the initial one on all interviewers. This would allow Field Office Supervisors (FOSs) and Crew Leaders (CLs) to check that interviewers continued to follow procedures as the survey finished.

2.4 Overall Flow of a Case within 2010 Person Interview

2.4.1 Sample Selection and Delivery

The CCM primary sampling unit is a block cluster, which consists of one or more geographically contiguous census blocks. Block clusters were formed to balance statistical and operational efficiencies. Two samples were selected to measure census coverage of housing units and the household population: the population sample (P sample) and the enumeration sample (E sample). The P sample is a sample of housing units and the persons within those housing units for the sample of block clusters selected for CCM operations. The E sample is a sample of census housing units and persons in the same block clusters as the P sample. The CCM sampling comprises a number of distinct processes from forming block clusters, selecting sample block clusters, implementing the sample reduction that reduced the number of sample block clusters, and eventually selecting addresses for the P and E samples.

A stratified sample of block clusters was selected for each state or state equivalent. An independent address list was created for each CCM sample block cluster. However, before the selection of the P and E samples, whole block clusters previously selected for CCM were dropped as part of the sample reduction associated with the nonsampling error reduction initiative. (See Section 2.3) From the remaining CCM block clusters, the P and E Sample addresses were selected. The selection of CCM housing units was dependent on the size of the block clusters. For block clusters with fewer than 80 CCM housing units, all units (the complete block cluster) were selected with certainty. For block clusters with 80 housing units or more, a subsample of units was selected. In addition, a sample of census housing units listed in the Census in the CCM clusters that were not listed by the CCM (based on the Initial Housing Unit Clerical Matching results) were also included in the PI workload. We refer to these units as Census-only cases. Once the cases were selected, they were sent to TMO and loaded into the CMOCS. For more on sampling of CCM cases and the sample design, see Census Coverage

Measurement Memo Series C-13-R1, 2010 Census Coverage Measurement: Sample Design—Revised (Fenstermaker, 2010).

2.4.2 Management of Cases in the Regional Census Centers

Once the cases were loaded into CMOCS, each RCC had their entire workload. Each RCC covered a section of the United States and had its own workload. The workload was split among teams that cover certain geographical areas within the RCC area. Each team was in charge of managing the cases and staff within the geographic area they covered. The main tool for this was CMOCS, based on the Regional Office Survey Control system, (the control system used for current demographic surveys). This system managed all the cases by allowing all of the following main functions:

- Setting up team geography – Each team covered a certain subset of geography and was responsible for all workload within that area. There were three or four teams within each region.
- Assigning field staff to teams and crews.
- Making initial assignments of all cases.
- Monitoring of cases from the field by displaying the most up-to-date outcome for each case.
- Reviewing cases as they were completed. Cases that were complete interviews were automatically sent to Headquarters but could still be reviewed by office staff. All other outcomes were shown on a screen for resolution where the office staff decided whether to accept that outcome or to reassign the case to another interviewer or the same interviewer to attempt another interview.
- Making reassignment of cases in the field – As cases were being worked, cases could be reassigned to different interviewers or could be reassigned to multiple interviewers if necessary.
- Reviewing multiple versions of a case – CMOCS alerts staff when there is more than one interview version of a case and the staff can then decide which version should be kept and which should be eliminated.
- Creating reports for monitoring both staff and case workload.
- Closing out the survey – CMOCS allowed the RCCs to close out the survey once all cases were received and reviewed. The office could do this at the Local Census Office level to better flow closeout and allow post production to begin on a flow instead of waiting for the entire survey to close out.

CMOCS handled both PI and PI RI as separate surveys.

2.4.3 Training

In general, a cascaded training process was used for CCM PI. This means that the highest level staff (RCC Staff) was given training first, who in turn then trained the next level of staff (FOSs, CLs, and Crew Leader Assistants (CLAs)), who then in turn trained the interviewers. This allowed everyone to be familiar with all the training and allowed better distribution of staff. For interviewer training, some of the RCCs also chose to do a staggered training (training on two different weeks) to better allow the same staff to train more people for consistency and best use of resources. Each interviewer training session required at least two trainers, but three were recommended. The third person was often office personnel observing and participating when

necessary. Master Trainers were also allowed to be used in lieu of cascaded training. RCCs could use this option if they had a skilled trainer available because they felt it led to a more consistent and better training.

The interviewer training duration was five days. The training manual was standardized to be read verbatim with custom regional topics being discussed on the last day. The training included both CCM concepts and hands-on training on the laptop, including the transmitting for the first time in order for trainees to pick up their first production assignments.

In addition, there was refresher training for interviewers that was held seven to ten days after they began their work. This was to review any questions, go over concepts they may have had trouble with, and to re-emphasize the most important concepts and anything people were not clear on now that they had some experience.

2.4.4 In the Field

Within each team, there were two to four FOSs who worked in the field. Each FOS would manage around four crews. Each crew was led by a cl and was recommended to have six interviewers. This structure ratio is less than originally planned, as it was one of the nonsampling error reduction initiatives enacted for 2010. The interviewers were instructed to meet with their crew leaders daily to go over payroll and to review the progress of their workload. This was also the time that an interviewer could ask any questions about individual cases. The CL reported to the FOS. The FOS also handled payroll and helped with any need the CL had including working with the RCC on procedural or technical problems.

Each interviewer had a Hewlett Packard Notebook laptop. It is the same laptop that current demographic surveys at the Census Bureau had just begun to use. It was loaded with case management software that allowed the interviewers to review their workload and launch the instrument for a case. Case management contained all the cases they currently had assigned and displayed the address and any notes or case history so far. It also allowed them to enter notes on a case. They could also check cases they completed to make sure they transmitted off their laptop. The transmissions from the laptop were supposed to be done each evening whether the interviewer had worked that day or not to make sure the case load was as current as possible. The transmissions were secured through a RSA token that used a random number generator and multiple passwords.

The instrument was programmed in BLAISE survey software and was a multi-path, multi-contact instrument. It allowed the interviewer to do the interview either by person or, when requested, by phone. It also allowed an interview at the sample address or with a proxy. It was programmed with the proxy rules and limited the ability to do a proxy until the right conditions were met. The instrument had several sections. The first section was to make contact and to determine the current status of the sample address. The second section was to collect a roster of all people living at the sample address at the time of interview, if it was an occupied housing unit. The next section determined if there were any inmovers since Census Day or individual outmovers. The following section collected the tenure and demographic information of the people rostered. The next section determined if each person had any alternate addresses. The following sections determined where each person lived and stayed most of the time around

Interview Day and around Census Day. For most occupied units, this is the end of the interview, but for cases where the house was occupied by a completely different set of people on Census Day (Whole Household of Outmovers), we repeat the process for those people. See Attachment A for a summary of each section and the main questions in the interview.

2.4.5 Post Processing

Cases were checked for completeness as they were received. They were then processed in waves during the operation to send to the next operation (Person Computer Geocoding and Clerical Residence Status Coding) to give as much time as possible for clerical matching without pushing the schedule back. Each case went through this post processing which put the data in a standardized format, cleaned off-path data, and set the residence status for each person listed in PI. The residence status assigned by computer in this operation is not the final status and was on the conservative side, marking all people as needing "Review" if there were any discrepancies or additional notes on the person.

3. Methodology

3.1 Questions to be Answered

The following is the list of questions that will be answered by this assessment. The focus of this study is to measure or gauge how efficient this operation was and to indicate how well the operation as a whole did collecting the information needed to make CCM a success.

3.1.1 How was the workload distributed over the length of the survey?

Using cost and progress data and review of when the output was received, we will look at the distribution of when cases were completed over the length of the survey. We will also discuss the initial sample.

3.1.2 How well did CCM Person Interview conform to the expected schedule and cost?

This section will review all aspects of the field budget estimates and review them against the actual. This will be done for all field positions and both PI and RI.

The cost results presented in this assessment were generated by program office staff using methods predating the U.S. Census Bureau's commitment to comply with the Government Accounting Office's cost estimating guidelines and the Society of Cost Estimating and Analysis best practices. Hence, while the Census Bureau believes these cost results are accurate and will meet the needs for which they will be used, the methods used for estimating costs of 2010 Census operations may not meet all of these guidelines and best practices. The Census Bureau will adhere to these guidelines in producing 2020 Census cost estimates.

3.1.3 How did Field staffing and training plans meet the needs for Person Interview production?

This section will review the planned staffing and training levels and compare them to the actual levels needed.

3.1.4 What were the Response Rate and Cooperation Rate for CCM Person Interview?

The response rate and cooperation rate let us know how successful CCM PI was in collecting data needed. The higher the response rate the more successful the operation. Response rate is defined as Completes / (All cases – Vacants - Not a Housing Units). Cooperation Rate is defined as (Completes + Vacant + Not a Housing Units) / All cases.

We will also study the reason cases were most commonly classified as a sufficient partial and also how much data we received for insufficient partial cases.

3.1.5 How much effort was necessary by interviewers to get each type of status for a case?

Using the tracking data from the instrument software, we will review the average number of count attempts and the average length of time within the instrument to get each status of a case (*Complete, Noninterview, Vacant, Not a Housing unit*). For *Complete* interviews, we will review the difference in these averages for a sample occupant respondent as compared to a proxy respondent to see how much more time is required to obtain a proxy respondent.

3.1.6 How effective was the automation of the CCM Person Interview Operation?

We will review all data problems that needed to be corrected in post processing and problems reported throughout the production stage, and how they needed to be corrected.

3.1.7 How often did Headquarters receive multiple versions of a specific case?

The CMOCS allowed the regions to send a different version of a case whenever they deemed it to be correct. We built in the ability for them to resend a case (either the same version or a different version) if they felt the data received in that version was more accurate than the previous instance submitted to Headquarters. This was to allow better control of the data within regions. We will review all versions received in output, how effective was the selection of the final production version through form selection, and how often version selection needed to be used.

3.1.8 How many production cases failed the quality check and needed to be replaced with Reinterview cases? What are the characteristics of those Reinterview cases?

Through the reinterview process, reinterview cases go through various steps to determine if a PI case was possibly falsified. At the end of the process, a reinterview case can indicate to the form selection software that the production case is suspect or definitely falsified. We will review how many cases were marked as suspect and how many times that led to failure and replacement of the PI case. We will also review the quality of those RI cases. The main output used is the "fail

file" created by PI RI Matching, Review, and Coding system (MaRCs)[6] output, RI output, and PI output.

3.1.9 How effective was Person Interview at rostering people?

This question will review how well PI did at collecting people as well as examine the type of people it collected. The following items will be included in this review:

- How many people were listed overall and by collection status (main roster, outmover, whole household of outmovers)
- Average household size overall, by type of respondent, and household relationship
- Sections where people were added (main roster, probes, review, outmover section)
- The overall demographics of the people collected
- Completeness of names collected

3.1.10 How effective was Person Interview at collecting alternate addresses?

An alternate address is any additional address a person rostered in PI and reported to have stayed at either around Census Day or in the last year. The person may or may not have been reported at the alternate address in the Census. This question will review how well PI did at collecting alternate addresses for people. The following items will be reviewed:

- Percentage of cases with alternate addresses
- Percentage of people with alternate addresses
- The section/type where the addresses were first reported
- The total number of addresses collected for each case by type of address
- The completeness of addresses provided
- How often an alternate address is shared by the people listed within the case
- How often the same address is listed for a person more than once

3.1.11 How did Person Interview do at collecting the timing of when people were living or staying at the Sample Address or Alternate Address?

We will review the information regarding time lived at the sample address in PI for each type of living situation to see how often the respondent could provide the specific timing information required to determine residency.

[6] PI RI MaRCs took the data from PI and PI RI and computer matched the person roster and housing unit status to determine if the information matched. If not, the system allowed a full review of the data for clerk either in NPC or the RCC to code the case for whether we suspected falsification.

3.1.12 How effective was Person Interview at collecting information to determine where people should be counted on Census Day?

While PI is not the final operation that determines Census Day residence status for CCM, this question will combine the results of collecting alternate address and dates of staying at all reported addresses to allow a determination of where people should be counted on Census Day based on PI information only. We will also review the post processing done at the end of PI data collection to see what residence status codes were sent onto the Person Matching operation.

3.1.13 What were the missing data rates for residence questions?

We will review the missing data rates for all the key residence questions including the inmover and outmover questions.

3.1.14 What impact does using proxies have on data quality?

We will review the distribution of sample occupant responses to determine if there is a difference in household size, the number of addresses listed, and missing data rates for interviews and those given by proxy respondents. We will also look at the ability to get complete information on proxies so that they could be contacted for the reinterview.

3.1.15 How often was the interview conducted in Spanish when Stateside? How often was the English interview used in Puerto Rico?

The interviewer marks the language in which they completed the interview at the end of each complete interview. We will use this variable to review how often the translations were used overall and by regions and states. We also report on feedback about the Spanish translation from lessons learned and field debriefings.

3.1.16 What were the differences and difficulties of conducting Person Interview in Puerto Rico?

PI interviews in Puerto Rico had a few challenges due to the differences from Stateside that were not just limited to Spanish being the primary language of respondents and interviewers. For example, the components and standards for an address in Puerto Rico are very different and much more complex than a Stateside address. In this section, we will review any difficulties that resulted from these differences, as well as changes that needed to be made to accommodate the different address system, both in materials and in the instrument.

3.1.17 How did components of the instrument work?

In this section, we will review more of the specific sections within the instrument to see how the wording and automation worked. These sections are of particular interest to PI because they involve checking the data quality or are used to improve coverage over Census operations.

3.1.18 What impact did the Nonsampling Error Reduction Initiatives have on the Person Interview operation?

As described in Section 2.3, the PI introduced some initiatives with the intent to improve the quality of the PI operation and its data. Many of the initiatives results cannot be measured quantitatively, but reports from observations and debriefings gave feedback on the impact from these initiatives.

3.2 Methods

All of the questions listed above will be answered through various outputs and observations. The following is a table of data available and how it will be used to answer the questions above.

Table 1: List of Data Sources Used in the Person Interview Assessment

Data Source	Purpose
2010 CCM PI Output	To perform tallies on data collected in PI by the interviewers.
2010 CCM PI RI Output	To perform tallies on data collected in PI RI for those cases that replaced PI.
2010 CCM PI Post Processing Output files	Selection of the analysis universe for final forms to be used. To review the number of versions received. Perform tallies on all post processing and residence status coding.
Trace Files of Keystrokes from PI and PI RI Automated Instruments	To review any data issues that may arise. To review the instrument for performance efficiency and interviewer/respondent comprehension. To review any post-interview correction for errors.
Reports for Coverage Measurement Operations Control System	To track field progress.
2010 PI Operation Cost and Progress data	To provide operational cost data to determine the total cost of the production and reinterview phases of the operation, as well as costs for mileage, hours, training, and overtime. To provide operational progress data to determine the total progress of the production and reinterview phases of the operation, as well as progress over time.
Master Activity Schedule – data	Comparison of actual start and finish dates to planned start and finish dates.
Staffing – tallies	Field staffing by position and overtime (authorized, invited to training, trained, worked, replacements) – both out in the field and in the field office.
Training data	To review the number and duration of training sessions and replacement training sessions.
Help Desk Information	To provide a high-level summary of the various help desk tickets.
Lesson Learned and Debriefings	Gathered from field and Headquarters staff involved with the planning and implementation of the project. To help qualitatively answer some of the assessment questions for both production and reinterview

Some other assumptions made when reporting the PI data are listed below:

1. All analysis will be done on the final form selected unless otherwise specified, including RI cases that replaced the production cases.
2. Puerto Rico will be included in all overall analyses except for alternate address collection. Puerto Rico will be reported separately in those sections and if otherwise stated.
3. Puerto Rico was managed from the Boston RCC. However when reporting at the RCC level, it will be reported separately from Boston.
4. Any errors identified during production processing were corrected and that data will be used in this assessment to provide the most accurate representation of the actual data results used for CCM processing[7].

4. Limitations

A. All PI data were reviewed during and after production for errors in the data. Due to the complexity of the interviews and some unexpected situations in the field, the PI instrument did not always code certain processing variables correctly. All known problems were logged and corrected in operations following PI when the parameters for the problem could be identified (See Footnote 6). We did the same corrections for this analysis. Naturally, this allows for the possibility of some human error in the corrections, although we verified all changes made. In addition, some problems did not have standard parameters that allowed corrections. Those problems are noted in the analysis where appropriate.

B. The BLAISE tracking data are connected to only one version of the case and only contains the interview attempt counts for that version. In situations when the RCC determined it was better to reassign a case as a new interview to another interviewer, any tracking data from the original attempts were lost. This resulted in some cases having very low count attempts or total time spent on the case. There is no way to identify the cases that lost this information.

C. If the instrument was not properly exited the timer within that case kept incrementing until either the case was entered again or the laptop was shutdown. This can lead to a few cases that are well over the expected average interview length.

5. Results

5.1 How was the workload distributed over the length of the survey?

The 2010 PI consisted of a subsample[8] of housing units from the original CCM listing sample (the P-Sample). In addition, we selected a sample of census listed housing units within the CCM clusters referred to as the E-Sample and identified in the housing unit matching phase the census-only cases to send to PI; that is housing units not listed by CCM but included in the

[7] Most of these problems were not fully documented or identified until after data had been forwarded to Person Matching. It was determined the best way to handle these cases was to report the problems and how to correct the problems directly to estimation staff that built the corrections into their processing.

[8] The sampling plan subsampled housing units within the clusters listed during initial listing based on cluster size. In addition, due to the decision to reduce the sample, whole clusters were dropped from CCM before performing PI sample selection. For more on sampling for PI, see the Operational Assessment on CCM Sampling.

E-Sample. Both types of these cases together make up the PI workload of 186,766 cases. See Table 2 for the distribution of the PI workload. As shown, only 4.32 percent (8,070 cases) were census-only cases compared to the 178,696 P-Sample cases. The workload was delivered in one delivery to all of the 12 RCCs at the beginning of the production cycle on June 14, 2010.

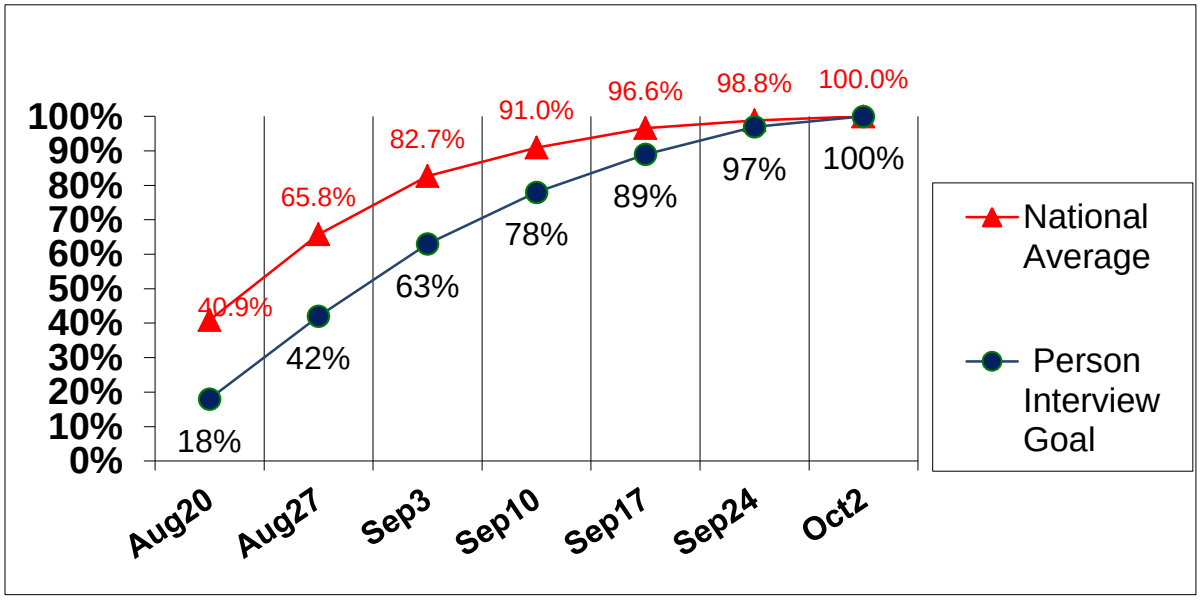
Table 2: Person Interview Workload by Type

	Count	Percent
P Sample cases	178,696	95.68
Census only cases	8,070	4.32
Total	186,766	100.00
Source: PI Sample File		

The official start date of CCM PI was August 14, 2010, but some of the production cases were assigned to FOSs and CLs at the end of their training session. The goal was for FOSs and CLs to work a few cases in order to be better acquainted with the procedures and issues involved in interviewing and have a better feel for how to supervise and guide their team of interviewers. Training was staggered with the FOSs trained before CLs, who were trained before the interviewers. FOS training began on July 6, 2010. FOS training was two weeks and at the end of the training session, they could begin doing live interviews. We received our first set of completed interviews on July 25, 2010. Some of the interviews were most likely done before July 25, but it was decided that was when the delivery of the data would begin. A total of 3,368 cases were completed as part of this live training cycle before the official beginning of the operation. CCM PI officially ended on October 2, but field work for RI lasted until October 14, 2010. Since some of the cases for PI were replaced by RI cases, the final case workload goes beyond October 2.

Before production began, Field Division set goals for each region to meet on a weekly basis. These were for the percent of cases to be completed by that date. Figure 1 displays the goals set and the national average for the percent of cases completed by each RCC and transmitted to Headquarters by that deadline.

Figure 1: Person Interview Workload Finish Rate Compared to Field Goals

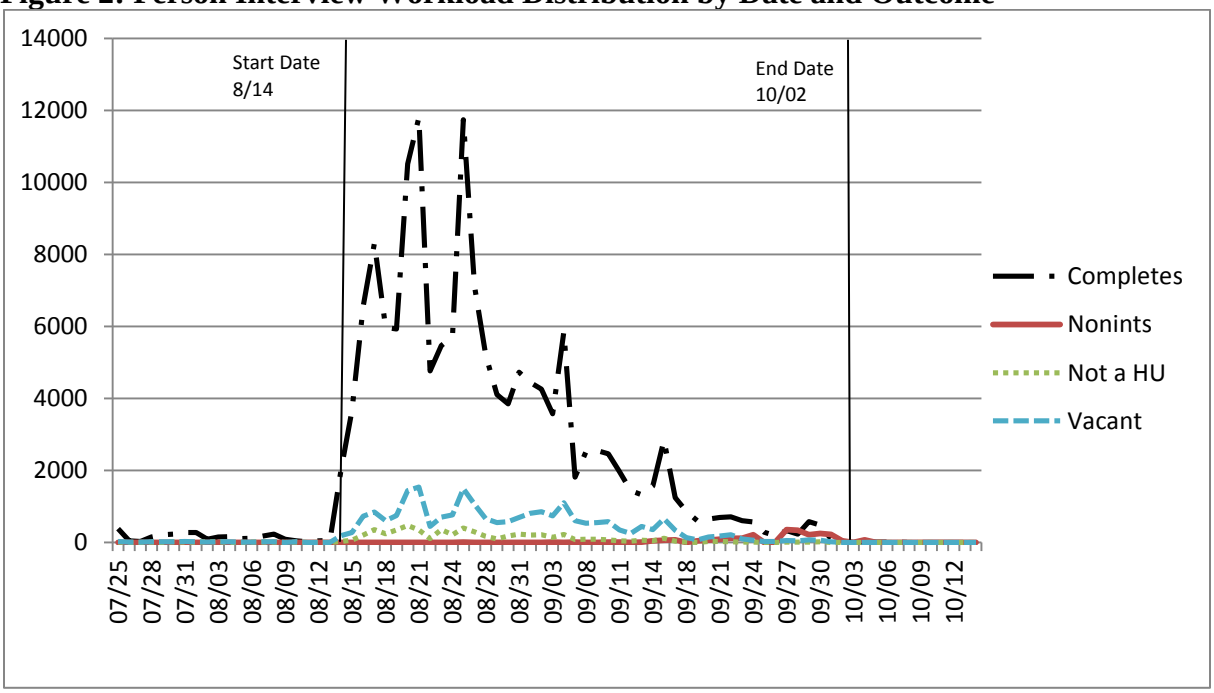


As you can see, PI exceeded the goal every week from the very beginning with almost half the cases done in the first week alone. Within the four week mark (September 10), 91.00 percent of cases had been completed and the last two weeks were used to complete the hard to contact cases.

The four graphs below depict the various outcome categories and present a good representation of when the cases were completed since each interviewer was asked to transmit every night, even if they did not work that day[9]. The instrument has an internal date collection as well, but it is dependent on the date set on the laptop and unfortunately some of the dates were not properly assigned on the laptop. Therefore, we consider the delivery date as a more accurate indication of the actual workload completion date. Figure 2 presents the overall distribution of all the PI cases by date received and outcome category (Complete, Noninterview, Not a Housing Unit, and Vacant). The dates marked on the figures are the planned/official start and finish dates.

[9] There is no way to track with certainty that interviewers always transmitted every night but it was expected and a mostly followed procedure that gives us a good overall picture of when work was done.

Figure 2: Person Interview Workload Distribution by Date and Outcome



The figure follows the pattern that was expected. In particular, the completion rate is very high after the official beginning (August 14, 2010) with most of the workload being resolved almost immediately and then trailing off as we got to the end of the data collection period. There are drops in the first three weeks but they seem to be around every six or seven days and may be caused by the interviewers reaching their 40 hour a week limit. It also shows that cases that are vacant or not a housing unit follow a very similar pattern, as they are very easy to resolve as well. Most noninterviews were received toward the end of the operation since field staff had no limit on the number of attempts to get the interview and could do multiple attempts all the way to the end of the survey.

Figure 3 is a subset of Figure 1, showing only the cases that did not result in an interview by date and outcome. This shows how the noninterviews came in at the very end of the survey as expected.

Figure 3: Person Interview Workload Distribution of Noninterviews by Date and Outcome

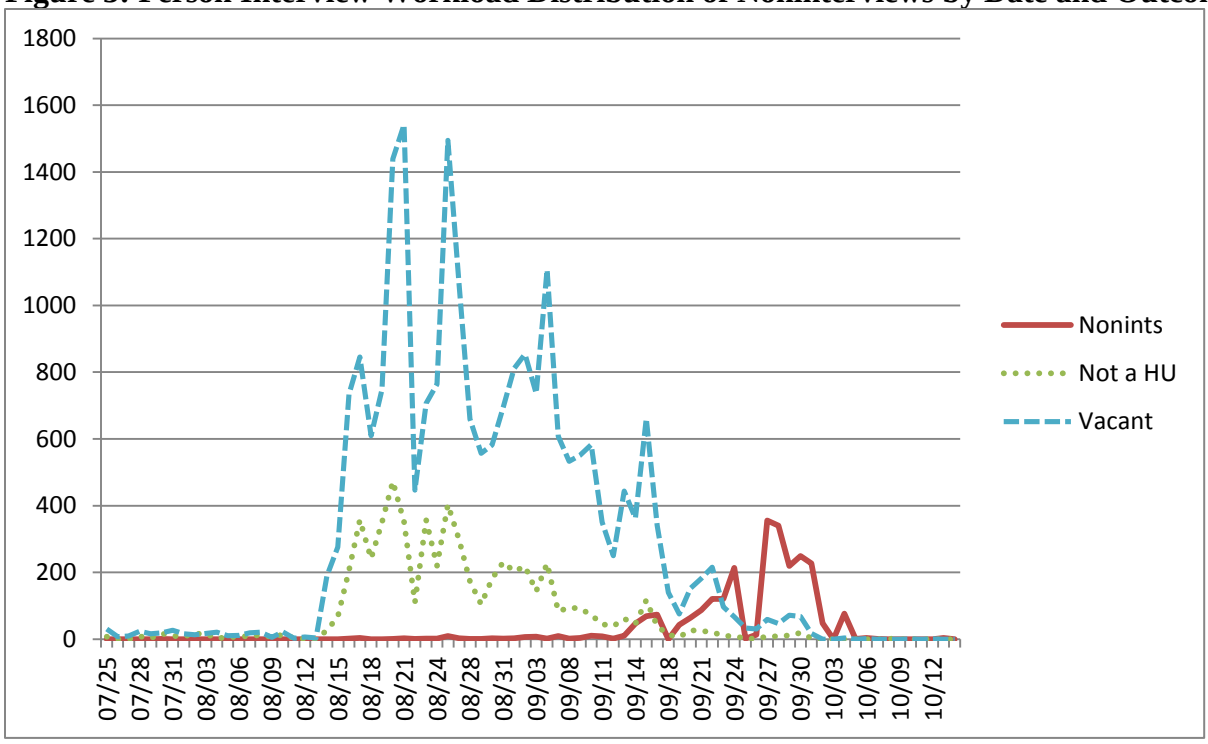
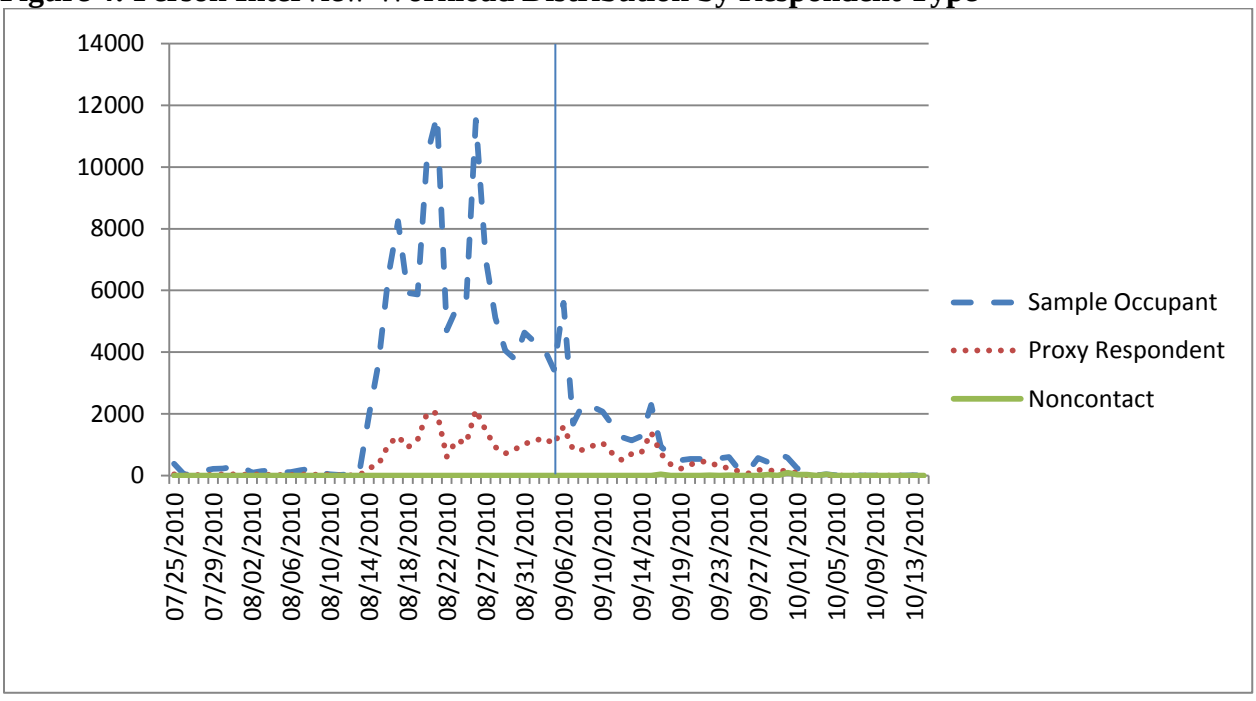


Figure 4: Person Interview Workload Distribution by Respondent Type



We wanted to limit proxy respondents as much as possible. Proxies were allowed immediately for units that were confirmed vacant or not housing units. Conversely, cases that were occupied units could only be worked by proxy immediately in situations where the cases were seasonal, all occupants were away until after October 2, or all occupants were incapacitated. In some other special situations, the CL could approve an early proxy for a case they deemed appropriate. Otherwise, the proxy could not be used unless an interviewer had attempted six visits and three weeks had elapsed since the first attempt. We would expect that there would be a spike in proxies in the beginning of the survey as interviewers resolve address vacant and not a housing unit cases. We expected another peak around September 6, which is the three-week mark from when most cases would have first been attempted (hence, when the computer would allow them to start attempting proxies for reluctant respondents). September 6 is marked on Figure 4 and 5. However, we did not get an initial spike of proxies, but a set of initial small peaks for proxies that do correspond to Figure 2, showing that vacant units and those that were not a housing unit were steadily resolved throughout the first few weeks.

Figure 5 displays the distribution of proxy respondents for just the complete interviews. In this case, we would expect to see a very low proxy count until September 6 and then a significant jump.

Figure 5: Person Interview Workload Distribution of Proxies for Complete Cases

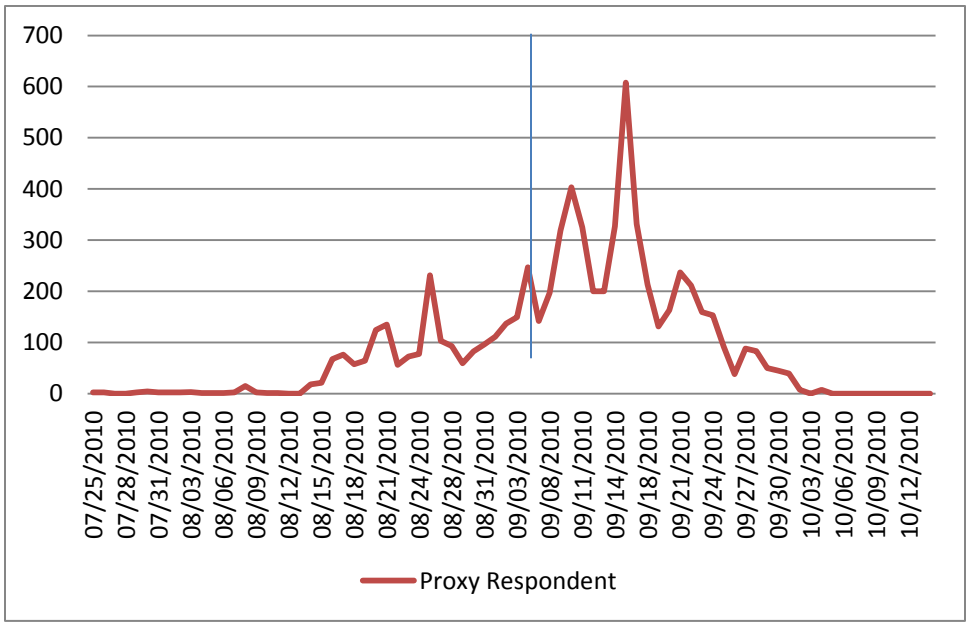
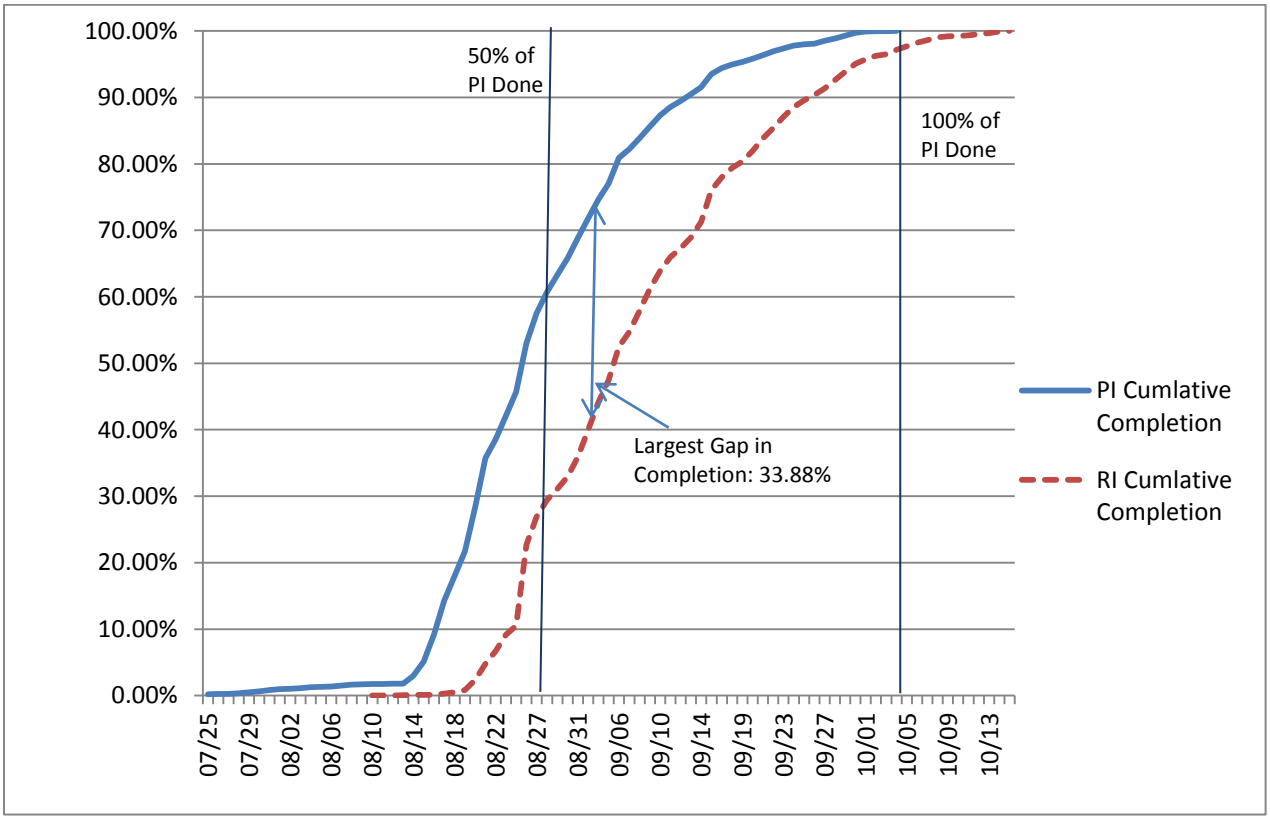


Figure 5 does show that the number of proxies increases greatly after September 6, but there are still a higher number than expected for proxy completes before September 6. This could result from the cases that were worked by the FOSs and CLs during the training session that would have had an earlier first attempt date. For more on why complete interviews used a proxy respondent, see Section 5.14.

It is also important that the workload distribution for RI cases follow a very similar completion rate as the PI workload with as few delays as possible. Figure 6 shows that the RI workload was completed in a similar rate to PI and that there was no backlog in sending RI cases for completion. The largest gap in the overall completion rate happened on August 31, with a gap of 32.88 percent. This is not a large gap given that RI began a couple of weeks later and was just getting into full production.

Figure 6: Rate of Completion for Person Interview and Person Reinterview



5.2 How well did Person Interview conform to the expected schedule and cost?

5.2.1 Schedule - How did actual start and completion dates compare to planned start and completion dates?

PI operations were planned from August 14, 2010 through October 2, 2010. As part of the Nonsampling Error Reduction Initiatives, FOSs and CLs were permitted to begin interviewing early as part of their hands-on training. PI began on July 14, 2010, and finished on schedule. PI RI was scheduled from August 19, 2010 through October 16, 2010. The operation started early on August 10, 2010, since cases were available, and finished on schedule. Table 3 shows the planned and actual dates when the field training was conducted.

Table 3: Person Interview and Person Reinterview Training Schedule

Training	Person Interview				Person Interview Reinterview			
	Planned		Actual		Planned		Actual	
	Start	End	Start	End	Start	End		
Field Operations Supervisors	7/9/2010	7/20/2010	7/6/2010	7/20/2010	7/9/2010	7/20/2010	7/9/2010	7/20/2010
Crew Leaders	7/26/2010	8/3/2010	7/19/2010	8/3/2010	7/26/2010	8/3/2010	7/19/2010	8/3/2010
Interviewers	8/9/2010	8/13/2010	8/9/2010	8/13/2010	8/9/2010	8/17/2010	8/9/2010	8/17/2010
Source: Decennial Master Activity Schedule								

There were 13 schedule Change Requests (CRs) to the Master Activity Schedule implemented for the PI operation. The CRs included date changes and logic corrections. Activity lines affected by the CRs were those of Cost and Progress, the Quality Profile, and PI RI MaRCS testing and production releases. Also affected were development of materials, geocoding, and training. There were no known issues or risks associated with implementing these CRs.

5.2.2 Costs – Was the operation over or under budget?

Overall, the PI operation was under budget. The Decennial Management Division (DMD) budget estimates assumed various factors. These assumptions were based on the results of prior field operations, as well as standardized and operation specific factors.

Assumptions included in the budget estimates, that were based on prior field operation results included:

- production rate per hour,
- field work hours per day,
- field work miles per day,
- training hours per day, and
- training miles per day.

In addition, standardized factors included salary, salary application rates, and mileage reimbursement rates. Operation specific factors included workload estimates and number of production days. Combining these factors as follows, the budget was estimated:

Total Cost = *Field Work Cost + Training Cost + Mileage Cost + Per Diem and Other Costs*

Field Work Cost is the cost of non-training wages and *Training Cost* is the cost of wages incurred during training hours, both excluding mileage. *Mileage Cost* is the total reimbursed mileage cost incurred during field work and training. *Per Diem and Other Costs* are the Meals and Incidental Expenses (M&IE), lodging cost, telephone costs and other expenses incurred during field work and training travel.

As you can see from the above equations, costs depend on many factors. These factors must be considered when comparing budget estimates to actual costs. For instance, when comparing training budget estimates to actual training costs, differences could be caused by either differences in the number of training staff, number of training days, training hours per day, salary rate, salary applications, or combinations of these. This section will attempt to explain why actual cost components varied from the budget estimate, whenever possible. In some instances, the data required to identify precise reasons for variation were not available or did not exist.

Table 4 provides the total budget and actual expenditures for the four components of total cost. A more detailed analysis of each component follows.

Table 4: Person Interview and Person Reinterview Total Cost by Component

Position	Budgeted Total Cost	Actual Total Cost	Difference of Budget to Actual Cost	Percent Over (+)/ Under Budget
Total Cost	**$32,073,356**	**$23,739,716**	**$8,333,640**	**25.98%**
Field Work Cost	$16,699,375	$13,033,473	$3,665,902	21.95%
Training Cost	$6,556,811	$4,089,437	$2,467,374	37.63%
Mileage Cost	$3,988,256	$5,390,083	($1,401,827)	(35.15%)
Per Diem and Other Costs	$4,828,910	$1,226,723	$3,602,187	74.60%

Source: Person Interview Cost & Progress Reports: Preliminary Total Cost ; Field Work; Training; Budget Totals Spreadsheet

*Data reflected are for the entire PI operation (both PI and PI RI combined).

5.2.2.1 Census Coverage Measurement Person Interview Cost Analysis

In this section, total cost is defined as all costs incurred during the operation. These costs, as defined in following sections, are field work cost, training cost, mileage cost, and Per Diem and other costs.

Table 5 provides the total budget and actual cost by position for both PI and PI RI.

Table 5: Person Interview and Person Reinterview Cost by Position

Position	Budgeted Total Cost	Actual Total Cost	Difference of Budget to Actual Cost	Percent Over(+)/ Under Budget
Total	$32,073,356	$23,739,716	**$8,333,640**	**25.98%**
Census Coverage Measurement Person Interview				
Subtotal-PI	**$25,681,998**	**$18,631,918**	**$7,050,080**	**27.45%**
Interviewer	$14,954,808	$10,771,266	$4,183,542	27.97%
Crew Leader Assistant	$3,314,659	$1,314,471	$2,000,188	60.34%
Crew Leader	$5,566,540	$4,633,014	$933,526	16.77%
Field Operations Supervisor	$1,845,991	$1,913,167	($67,176)	(3.64%)
Census Coverage Measurement Person Interview Reinterview				
Subtotal-PI RI	**$6,391,358**	**$5,107,797**	**$1,283,561**	**20.08%**
Reinterviewer	$3,296,306	$2,492,736	$803,570	24.38%
Crew Leader Assistant	$1,136,784	$443,263	$693,521	61.01%
Crew Leader	$1,183,244	$1,297,758	($114,514)	(9.68%)
Field Operations Supervisor	$775,024	$874,041	($99,017)	(12.78%)
Source: Person Interview Cost & Progress Reports: Preliminary Total Cost				

Total Cost Summary

Overall, the 2010 CCM PI operation was under budget by $8,333,640 (25.98 percent). PI was under budget by $7,050,080 (27.45 percent) and PI RI was under budget by $1,283,561 (20.08 percent).

Total Cost by Position

PI was under budget. Interviewer and Crew Leader Assistant (CLA) costs contributed most to being under budget. Interviewer cost was under budget by $4,183,542 (27.97 percent) and CLA cost was under budget by $2,000,188 (60.34 percent). CL cost was also under budget by $933,526 (16.77 percent). Though these positions were under budget, FOS cost was over budget by $67,176 (3.64 percent). However, this cost had minimal effect on the overall cost for PI.

PI RI cost was also under budget. Reinterviewer and CLA costs exhibited a similar cost ratio to that of PI. Reinterviewer cost was under budget by $803,570 (24.38 percent), while CLA cost was under budget $693,521 (61.01 percent). These lesser costs were offset by higher costs than budgeted for CLs and FOSs. PI RI CL cost was over budget by $114,514 (9.68 percent) and FOS cost was over budget by $99,017 (12.78 percent).

5.2.2.2 Cost Per Case

In this section, cost per case is defined as the total cost incurred for each case in the workload. Table 6 provides the budget and actual cost per case in the workload by position for both PI and PI RI.

Table 6: Person Interview and Reinterview Cost Per Case by Position

Position	Budgeted Total Cost	Actual Total Cost	Budgeted Total Cost per Case *	Actual Cost per Case †	Difference of Budget to Actual Cost per Case	Percent Over(+)/ Under Budget
Census Coverage Measurement Person Interview						
Subtotal-PI	**$25,681,998**	**$18,631,918**	**$125.00**	**$99.76**	**$25.23**	**20.19%**
Interviewer	$14,954,808	$10,771,266	$72.79	$57.67	$15.11	20.76%
Crew Leader Assistant	$3,314,659	$1,314,471	$16.13	$7.04	$9.09	56.37%
Crew Leader	$5,566,540	$4,633,014	$27.09	$24.81	$2.29	8.44%
Field Operations Supervisor	$1,845,991	$1,913,167	$8.98	$10.24	($1.26)	(14.01%)
Census Coverage Measurement Person Interview Reinterview						
Subtotal-PI RI	**$6,391,358**	**$5,107,797**	**$172.81**	**$166.13**	**$6.68**	**3.86%**
Reinterviewer	$3,296,306	$2,492,736	$89.13	$81.08	$8.05	9.03%
Crew Leader Assistant	$1,136,784	$443,263	$30.74	$14.42	$16.32	53.09%
Crew Leader	$1,183,244	$1,297,758	$31.99	$42.21	($10.22)	(31.94%)
Field Operations Supervisor	$775,024	$874,041	$20.96	$28.43	($7.47)	(35.66%)

Source: Person Interview Cost & Progress Reports: Preliminary Total Cost, PI Progress to Date, RI Progress to Date

*Budgeted Total Person Interview Workload was 205,464 and Budgeted Total Person Interview Reinterview Workload was 36,985.

†Actual Total Person Interview Workload was 186,766 and Actual Total Person Interview Reinterview Workload was 30,745 (90.90 percent and 83.13 percent of expected workloads, respectively.)

Cost Per Case Summary
PI actual cost per case was $99.76. This is $25.23 (20.19 percent) per case less than expected. The PI RI actual cost per case was $166.13. This is $6.68 (3.86 percent) per case less than expected.

Cost Per Case By Position
PI was more efficient than expected. Interviewers and CLAs greatly contributed to this efficiency. Interviewer cost was under budget by $15.11 (20.76 percent) per case and CLA cost was under budget by $9.09 (56.37 percent) per case. CL cost was also under budget by $2.29 (8.44 percent) per case; however, FOS cost was over budget and less efficient by $1.26 (14.01 percent) per case.

PI RI was slightly more efficient than expected. RI cost was under budget by $8.05 per case (9.03 percent) and CLA cost was quite under budget by $16.32 per case (53.09 percent). Though PI RI was more efficient overall per case, PI RI CLs and FOSs were much less efficient. CL cost was over budget by $10.22 per case (31.94 percent) and FOS cost was over budget by $7.47 per case (35.66 percent).

5.2.2.3 Field Work Costs

In this section, field work cost is defined as the cost of non-training wages. For the purpose of this section, mileage costs are not included; however, they are discussed in Section 5.2.2.5.

Table 7 provides the budget and actual field work costs by position for both PI and PI RI.

Table 7: Person Interview and Reinterview Field Work Cost by Position

Position	Budgeted Field Work Hours Cost	Actual Field Work Hours Cost	Difference of Budget to Actual Cost	Percent Over(+)/ Under Budget
Total	**$16,699,375**	**$13,033,473**	**$3,665,902**	**21.95%**
Census Coverage Measurement Person Interview				
Subtotal-PI	**$13,685,905**	**$10,325,198**	**$3,360,707**	**24.56%**
Interviewer	$7,376,698	$5,668,352	$1,708,346	23.16%
Crew Leader Assistant*	$2,278,341	$730,585	$1,547,756	67.93%
Crew Leader	$3,051,898	$2,785,432	$266,466	8.73%
Field Operations Supervisor	$978,968	$1,140,830	($161,862)	(16.53%)
Census Coverage Measurement Person Interview Reinterview				
Subtotal-PI RI	**$3,013,470**	**$2,708,274**	**$305,196**	**10.13%**
Reinterviewer	$1,511,599	$1,240,488	$271,111	17.94%
Crew Leader Assistant*	$507,839	$228,714	$279,125	54.96%
Crew Leader	$607,772	$726,189	($118,417)	(19.48%)
Field Operations Supervisor	$386,260	$512,884	($126,624)	(32.78%)
Source: Person Interview Cost & Progress Reports: Current Employee Cost – Field Work				

*Crew Leaders Assistants were trained as Interviewers.

Field Work Cost Summary

Overall, the cost for field work associated with the 2010 CCM PI operation was under budget by $3,665,902 (21.95 percent). PI fieldwork cost was under budget by $3,360,707 (24.56 percent) and PI RI fieldwork cost was under budget by $305,196 (10.13 percent).

Field Work Cost By Position

Lower PI Interviewer and CLA costs greatly contributed to PI being completed under budget. Interviewer field work cost was under budget by $1,708,346 (23.16 percent) and CLA field work cost was also significantly under budget by $1,547,756 (67.93 percent). CL fieldwork cost was also under budget by $266,466 (8.73 percent); however, FOS fieldwork cost was over budget by $161,862 (16.53 percent). PI RI fieldwork cost was also under budget. Reinterviewer fieldwork cost was under budget by $271,111 (17.94 percent) and CLA fieldwork cost was under budget

by $279,125 (54.96 percent). However, CL and FOS fieldwork cost was over budget. CL fieldwork cost was over budget by $118,417 (19.48 percent) and FOS fieldwork cost was over budget by $126,624 (32.78 percent).

5.2.2.4 Training Costs

In this section, training cost is defined as the cost of wages incurred during training. For the purpose of this section, mileage costs are not included; however, they are discussed in Section 5.2.2.5.

Table 8 provides the budget and actual training cost by position for both PI and PI RI.

Table 8: Person Interview and Reinterview Training Cost by Position

Position	Budgeted Training Hours Cost	Actual Training Hours Cost	Difference of Budget to Actual Cost	Percent Over(+)/Under Budget
Total	**$6,556,811**	**$4,089,437**	**$2,467,374**	**37.63%**
Census Coverage Measurement Person Interview				
Subtotal-PI	**$4,956,521**	**$3,310,570**	**$1,645,951**	**33.21%**
Interviewer	$3,465,288	$2,275,977	$1,189,311	34.32%
Crew Leader Assistant*	$465,416	$208,821	$256,595	55.13%
Crew Leader	$783,799	$619,002	$164,797	21.03%
Field Operations Supervisor	$242,018	$206,770	$35,248	14.56%
Census Coverage Measurement Person Interview Reinterview				
Subtotal-PI RI	**$1,600,290**	**$778,867**	**$821,423**	**51.33%**
Reinterviewer	$861,288	$469,896	$391,392	45.44%
Crew Leader Assistant*	$448,707	$57,211	$391,496	87.25%
Crew Leader	$180,726	$158,939	$21,787	12.06%
Field Operations Supervisor	$109,569	$92,821	$16,748	15.29%
Source: Person Interview Cost & Progress Reports: Current Employee Cost – Training				

*Crew Leaders Assistants were trained as Interviewers.

Training Cost Summary

Overall, the cost for training associated with the 2010 CCM PI operation was under budget by $2,467,374 (37.63 percent). PI training cost was under budget by $1,645,951 (33.21 percent) and PI RI training cost was under budget by $821,423 (51.33 percent). This is mostly because we did not train to full staffing allotments.

Training Cost By Position

PI training cost was under budget. Interviewer training cost was under budget by $1,189,311 (34.32 percent) and CLA training cost was very much under budget by $256,595 (55.13 percent). CL training cost was under budget by 164,797 (21.03 percent) and FOS training cost was under budget by $35,248 (14.56 percent).

PI RI training cost was very much under budget. Reinterviewer training cost was under budget by $391,392 (45.44 percent) and CLA training cost was extremely under budget by $391,496 (87.25 percent). Unlike field work cost trends, CL and FOS training costs were also under budget. CL training cost was under budget by $21,787 (12.06 percent) and FOS training cost was under budget by $16,748 (15.29 percent).

5.2.2.5 Mileage Costs

In this section, mileage costs are defined as the total reimbursed mileage costs incurred for field work and training. During PI, field staff were reimbursed at a rate of $0.50 per mile.

Table 9 provides the budget and actual mileage costs by position for both PI and PI RI.

Table 9: Person Interview and Reinterview Mileage Cost by Position

Position	Budgeted Miles Cost**	Actual Miles Cost**	Difference of Budget to Actual Cost	Percent Over(+)/Under Budget
Total	**$3,988,256**	**$5,390,083**	**($1,401,827)**	**(35.15%)**
Census Coverage Measurement Person Interview				
Subtotal-PI	**$3,245,000**	**$4,104,234**	**($859,234)**	**(26.48%)**
Interviewer	$1,769,319	$2,449,315	($679,996)	(38.43%)
Crew Leader Assistant*	$570,902	$287,548	$283,354	49.63%
Crew Leader	$707,334	$978,535	($271,201)	(38.34%)
Field Operations Supervisor	$197,445	$388,837	($191,392)	(96.93%)
Census Coverage Measurement Person Interview Reinterview				
Subtotal-PI RI	**$743,256**	**$1,285,849**	**($542,593)**	**(73.00%)**
Reinterviewer	$337,521	$662,443	($324,922)	(96.27%)
Crew Leader Assistant*	$180,238	$123,993	$56,245	31.21%
Crew Leader	$145,325	$307,739	($162,414)	(111.76%)
Field Operations Supervisor	$80,172	$191,676	($111,504)	(139.08%)
Source: Person Interview Cost & Progress Reports: Current Employee Cost – Field Work				

*Crew Leaders Assistants were trained as Interviewers.

**Mileage reflects miles cost for both field work and training.

Mileage Cost Summary

Overall, the mileage cost for the 2010 CCM PI operation was over budget by $1,401,827 (35.15 percent). Mileage cost for PI was over budget by $859,234 (26.48 percent) and mileage cost for PI RI was also over budget by $542,593 (73.00 percent).

Mileage Cost By Position

Though total PI cost was under budget, mileage cost for PI was over budget. Mileage cost for Interviewers was over budget by $679,996 (38.43 percent), greatly contributing to the higher mileage expenditures, while mileage cost for CLAs was under budget by $283,354 (49.63 percent). Also contributing to the higher expenditures were mileage costs for CLs and FOSs. Mileage cost for CLs was over budget by $271,201 (38.34 percent) and mileage cost for FOSs was over budget by $191,392 (96.93 percent).

Mileage cost for PI RI was also over budget, exhibiting the same cost trends as PI. Mileage cost for Reinterviewers was over budget by $324,922 (96.27 percent), while mileage cost for CLAs was under budget by $56,245 (31.21 percent). Mileage cost for CLs was very over budget by $162,414 (111.76 percent) and mileage cost for FOSs was considerably over budget by $111,504 (139.08 percent).

5.2.2.6 Per Diem and Other Costs

In this section, Per Diem and other costs are defined as the M&IE, lodging cost, telephone costs, and other expenses incurred during field work and training. For the purpose of this section, mileage costs are not included.

Table 10 provides the budget and actual Per Diem and other costs by position for both PI and PI RI.

Table 10: Person Interview Operation Per Diem and Other Costs by Position

Position	Budgeted Per Diem Reimbursement	Actual Per Diem Reimbursement***	Difference of Budget to Actual Cost	Percent Over(+)/ Under Budget
Total	**$4,828,910**	**$1,226,723**	**$3,602,187**	**74.60%**
Census Coverage Measurement Person Interview				
Subtotal-PI	****$3,794,570**	**$891,916**	**$2,902,654**	**76.49%**
Interviewer	$2,343,500	$377,622	$1,965,878	83.89%
Crew Leader Assistant*	$0	$87,517	($87,517)	Not Applicable
Crew Leader	$1,023,510	$250,045	$773,465	75.57%
Field Operations Supervisor	$427,560	$176,730	$250,830	58.67%
Census Coverage Measurement Person Interview Reinterview				
Subtotal-PI RI	**$1,034,340**	**$334,807**	**$699,533**	**67.63%**
Reinterviewer	$585,900	$119,909	$465,991	79.53%
Crew Leader Assistant*	$0	$33,345	($33,345)	Not Applicable
Crew Leader	$249,420	$104,891	$144,529	57.95%
Field Operations Supervisor	$199,020	$76,660	$122,360	61.48%

Source: Person Interview Cost &Progress Reports: Current Employee Cost – Training; Current Employee Cost – Field Work; Preliminary Total Cost

*Crew Leaders Assistants were trained as Interviewers.

**Telephone budget rolled up into PI Production Interviewers.

*** Per Diem reflects Per Diem costs for both field work and training.

Per Diem Costs Summary

Overall, the 2010 CCM PI operation Per Diem costs were under budget by $3,602,187 (74.60 percent). PI Per Diem costs were under budget by $2,902,654 (76.49 percent) and PI RI Per Diem costs were under budget by $699,533 (67.63 percent).

Per Diem Costs By Position

Per Diem costs for PI were also under budget for all positions except CLAs. Per Diem costs for Interviewers were under budget by $1,965,878 (83.89 percent), Per Diem costs for CLs were under budget by $773,465 (75.57 percent), and Per Diem costs for FOSs were under budget by $250,830 (58.67 percent). Since no budget was allocated for CLAs, the cost was over budget by $87,517. Since Per Diem was under budget and mileage was over, this tells us that CLs and FOSs had to do much more driving within the area they worked than expected. This might be attributable to the fact that they had to pick up payroll almost daily and they traveled to the interviewers instead of coordinating interviewers coming to them. With an automated payroll in the future, the associated mileage cost should decrease significantly.

Per Diem costs for PI RI were also very much under budget, reflecting the same trends as PI. Per Diem costs for Reinterviewers were under budget by $465,991 (79.53 percent), Per Diem costs for CLs were under budget by $144,529 (57.95 percent), and Per Diem costs for FOSs were under budget by $122,360 (61.48 percent). Again, since no budget was allocated for CLAs, the cost was over budget by $33,345.

5.2.3 Productivity Rates

This section analyzes the effort required to complete a single unit of work in terms of field work (non-training) hours and mileage charged.

5.2.3.1 Production Rates for Person Interview and Person Reinterview

In this section, fieldwork rate is defined as the effort required to complete a single unit of work in terms of fieldwork (non-training) hours.

Table 11 provides the budget and actual field work production rates by position for both PI and PI RI.

Table 11: Person Interview and Reinterview Production Rate by Position

Position	Budgeted Field Work Hours	Actual Field Work Hours	Budgeted Cases per Production Hour*	Actual Cases per Production Hour†	Difference of Budget to Actual Cases per Production Hour	Percent More(+)/Less Efficient
Census Coverage Measurement Person Interview						
Total-Production	**818,859**	**612,631**	**0.25**	**0.30**	**0.05**	**21.50%**
Interviewer	456,492	354,666	0.45	0.53	0.08	17.00%
Crew Leader Assistant	140,625	44,599	1.46	4.19	2.73	186.62%
Crew Leader	171,327	155,893	1.20	1.20	0.00	0.10%
Field Operations Supervisor	50,415	57,473	4.08	3.25	(0.83)	(20.26%)
Census Coverage Measurement Person Interview Reinterview						
Total-Reinterview	**179,126**	**157,642**	**0.21**	**0.20**	**(0.01)**	**(5.54%)**
Reinterviewer	93,884	76,675	0.39	0.40	0.01	1.79%
Crew Leader Assistant	31,070	14,360	1.19	2.14	0.95	79.86%
Crew Leader	34,225	40,770	1.08	0.75	(0.33)	(30.22%)
Field Operations Supervisor	19,947	25,838	1.85	1.19	(0.66)	(35.82%)

Source: Person Interview Cost & Progress Report: Current Employee Cost – Field Work

*Budgeted Total Person Interview Workload was 205,464 and Budgeted Total Person Interview Reinterview Workload was 36,985.

†Actual Total Person Interview Workload was 186,766 and Actual Total Person Interview Reinterview Workload was 30,745 (90.90 percent and 83.13 percent less than expected, respectively.)

Production Rate Summary

During PI, 0.30 cases were completed per hour. This is 0.05 more cases per hour (21.50 percent more efficient) than expected. During PI RI, 0.20 cases were completed per hour. This is 0.01 fewer cases per hour (5.54 percent less efficient) than expected.

Production Rate By Position

The production rate for PI was more efficient than expected. Production rates for Interviewers and especially CLAs were more efficient than planned by 0.08 more cases per hour (17.00 percent more efficient) and 2.73 more cases per hour (186.62 percent more efficient), respectively. Though production rates for Interviewers and CLAs were more efficient, the production rate for CLs was as expected at 1.20 cases per hour, while the production rate for FOSs was less efficient by 0.83 cases per hour (20.26 percent less efficient).

Though the production rate for PI was more efficient than expected, the production rate for PI RI was less efficient. The production rates for Reinterviewers and CLAs were more efficient by 0.01 more cases per hour (1.79 percent more efficient) and 0.95 more cases per hour (79.86 percent more efficient), respectively. However, this efficiency was offset by the less efficient production rates of CLs and FOSs. The production rates for CLs and FOSs were less efficient by 0.33 fewer cases per hour (30.22 percent less efficient) and 0.66 fewer cases per hour (35.82 percent less efficient), respectively.

5.2.3.2 Mileage Rates for Person Interview Operations

In this section, mileage rate is defined as the number of miles charged to complete a unit of work.

Table 12 provides the budgeted and actual mileage rates by position for both PI and PI RI.

Table 12: Person Interview and Reinterview Mileage Rate by Position

Position	Budgeted Total Miles	Actual Total Miles	Budgeted Total Miles / Case *	Actual Miles / Case†	Difference of Budget to Actual Miles per Case	Percent More(+)/ Less(-)
Census Coverage Measurement Person Interview						
Total-PI	**6,489,998**	**8,208,468**	**31.59**	**43.95**	**(12.36)**	**(39.14%)**
Interviewer	3,538,643	4,898,630	17.22	26.23	(9.01)	(52.29%)
Crew Leader Assistant	1,141,794	575,095	5.56	3.08	2.48	44.59%
Crew Leader	1,414,669	1,957,070	6.89	10.48	(3.59)	(52.19%)
Field Operations Supervisor	394,892	777,673	1.92	4.16	(2.24)	(116.65%)
Census Coverage Measurement Person Interview Reinterview						
Total-PI RI	**1,486,490**	**2,571,698**	**40.19**	**83.65**	**(43.45)**	**(108.12%)**
Reinterviewer	675,043	1,324,885	18.25	43.09	(24.84)	(136.10%)
Crew Leader Assistant	360,463	247,985	9.75	8.07	1.68	17.24%
Crew Leader	290,642	615,477	7.86	20.02	(12.16)	(154.74%)
Field Operations Supervisor	160,342	383,351	4.34	12.47	(8.13)	(187.61%)

Source: Person Interview C&P Report: Current Employee Cost – Field Work

*Budgeted Total Person Interview Workload was 205,464 and Budgeted Total Person Interview Reinterview Workload was 36,985.

†Actual Total Person Interview Workload was 186,766 and Actual Total Person Interview Reinterview Workload was 30,745 (90.90 percent and 83.13 percent less than expected, respectively).

Mileage Rate Summary

During PI, 43.95 miles were charged overall per case. This is 12.36 more miles per case (39.14 percent) than expected. During PI RI, an overwhelming 83.65 miles were charged per case. This is 43.45 more miles per case (108.12 percent) than expected.

Mileage Rate By Position

The mileage rate for PI was higher than expected. The mileage rate for Interviewers was higher by 9.01 miles per case (52.29 percent), while the mileage rate for CLAs was lower by 2.48 miles per case (44.59 percent). Mileage rates for CLs and FOSs were higher by 3.59 miles per case (52.19 percent) and 2.24 miles per case (116.65 percent), respectively.

The mileage rate for PI RI was also higher than expected, with the same trends as PI. The mileage rate for Reinterviewers was higher by 24.84 miles per case (136.10 percent), while the mileage rate for CLAs was lower by 1.68 miles per case (17.24 percent). Mileage rates for CLs and FOSs were higher by 12.16 miles per case (154.74 percent) and 8.13 miles per case (187.61 percent), respectively.

5.3 How did Field staffing and training plans meet the needs for Person Interview production?

The Field Division provided a staffing authorization to each RCC. This authorization provided an upper limit for hiring in each RCC. RCC staff hired for each position at their discretion based on their regional implementation plans for the PI operations.

Table 13 shows the staffing authorized and trained for PI production and PI RI, by field position.

Table 13: Person Interview and Reinterview Field Staffing

	Production Staff				Reinterview Staff			
	Lister / Interviewer	Crew Leader Assistant	Crew Leader	Field Office Supervisor	Reinter-viewer	Crew Leader Assistant	Crew Leader	Field Office Supervisor
Person Interview Staff Authorized	4,291	536	659	167	709	111	136	74
Person Interview Staff Trained	3,758	323	584	156	678	78	137	73
Difference	-533	-213	-75	-11	-31	-33	1	-1

Source: Data provided by FLD

Authorized staffing levels were more than sufficient to perform and complete both PI production and PI RI with PI hiring only 85.28 percent of their authorized staff and RI hiring only 93.79 percent of their authorized staff. In addition, there was little sign of attrition with only 9.54 percent (460) of PI staff being trained after the operation started and only 3.31 percent (32) of RI staff being trained after the RI operation started.

5.4 What were the Response Rate and Cooperation Rate for Person Interview?

In this section, we will review the final outcome of the cases as well as the response rate and cooperation rate for the final versions of each PI case. For this assessment, response rate is

defined as Completes / (All cases – Vacants - Not a Housing Units). Cooperation Rate is defined as (Completes + Vacants + Not a Housing Units) / All cases.

For this report, the PI outcomes are grouped into outcome types. Completes are all interviews that were completed. This includes sufficient partials, which are cases where all the questions were asked but the respondent answered "Don't Know" or "Refused" on at least one of the key questions needed to determine residence status. Completes also include cases where the RI case that replaced the PI case did not do a full interview[10]. Vacants are any addresses that are housing units and were vacant on Interview Day. Not a Housing Unit refers to addresses that were determined to not be a housing unit on Interview Day (e.g., under construction, turned into a business). Noninterviews consist of refusals, break off interviews, unable to locate and unable to find a knowledgeable respondent; as well as other situations.

Table 14 provides the counts for each outcome the instrument could assign, grouped into their outcome types.

[10] During the PI RI interview a respondent could indicate that they did conduct a PI interview. In this case, the PI RI interview only collects a roster and then is considered complete. In the matching of the rosters to PI, it can be found that the PI version of the case is incorrect and the PI RI case should be the one used. In this situation, all we have is the Roster and no other demographic or residency information, but in the field this was considered a complete case. See Section 0.

Table 14: Person Interview Outcomes

Outcome Type/ Category	Outcome: Definition	Count	Percent
Completes		**155,243**	**83.12**
	201: Complete Interview with Sample Respondent	140,330	75.14
	202: Complete RI case (full PI not required)	244	0.13
	203: Sufficient Partial with Sample Respondent	7,783	4.17
	208: Complete Interview with Proxy Respondent	3,270	1.75
	209: Sufficient Partial with Proxy Respondent	3,616	1.94
Noninterview		**2,477**	**1.33**
	204: Noninterview in RI case (where a full PI should have been conducted but the interview was broken off)	6	0.00
	213: Language Problem	11	0.01
	216: No knowledgeable respondent available	150	0.08
	218: Refusal	1,243	0.67
	219: Other Noninterview	1,067	0.57
Vacant		**23,106**	**12.37**
	326: Vacant confirmed by typical proxy respondent (e.g. neighbor)	19,887	10.65
	327: Vacant by non-typical proxy type or observation	3,219	1.72
Not a Housing Unit		**5,940**	**3.18**
	333: Not a housing unit confirmed by a typical proxy respondent	4,631	2.48
	334: Not a housing unit by non-typical proxy or observation	1,309	0.70
Total		**186,766**	**100.00**
Source: Person Interview and Reinterview Output			

Overall, of the 186,766 cases, 83.12 percent (155,243) of cases were classified as a Complete. Of the rest, 15.55 percent (29,046 cases) were either vacant or not a housing unit. Only 1.33 percent (2,477) cases ended up being noninterviews.

Table 15 shows the response rates and the cooperation rate for CCM PI by RCC. Puerto Rico was managed by the Boston RCC, but for this report Puerto Rico will be treated as its own office. Overall, the final Response Rate was 98.43 percent and the final cooperation rate was 98.68 percent.

Table 15: Person Interview Response and Cooperation Rates by RCC (plus Puerto Rico)

Regional Census Center	Workload	Response Rate (Completes/Eligible Cases)	Cooperation Rates (Completes+ Vacants+Not a Housing unit/ All Cases)
Boston	13,513	97.04	97.46
New York	8,703	95.44	95.94
Philadelphia	13,645	98.02	98.31
Detroit	12,299	99.22	99.36
Chicago	12,818	97.87	98.17
Kansas City	12,229	99.13	99.27
Seattle	15,393	99.13	99.23
Charlotte	16,730	98.76	98.95
Atlanta	16,864	97.71	98.15
Dallas	16,423	99.15	99.29
Denver	23,049	99.00	99.18
Los Angeles	17,099	98.63	98.78
Puerto Rico	8,001	99.53	99.63
Overall	**186,766**	**98.43**	**98.68**
Source: Person Interview and Reinterview Output			

The response rate for all RCCs was 97.04 percent or higher except for the New York RCC at 95.44 percent. Five RCCs had response rates of 99.00 percent or higher. The Puerto Rico Office had the highest response rate at 99.53 percent. All RCCs followed a similar pattern for their cooperation rates.

The response rate for the personal visit phase of the 2000 A.C.E PI was 99.8 percent, higher than the 2010 overall response rate of 98.43 percent. A comparison of 2000 to 2010 shows that there were many factors in the lower response rate for 2010. See "Comparisons and Explanations of Differences in Person Interview Noninterview Rates from the 2000 Accuracy and Coverage Evaluation and the 2010 Census Coverage Measurement Programs" in the references for a full review of all the findings. Below are the five procedural changes from 2000 that may have had an effect on the 2010 response rate.

1. There was a change in the proxy procedures from 2000 to 2010. In 2000, if a respondent refused, the interviewer was allowed to look for a proxy respondent. The rules in 2010 were more stringent and did not allow for proxy in the case when a sample address occupant refused to participate. In addition, the general proxy rules were programmed into the 2010 instrument such that it would not allow an early proxy unless the interviewer provided a valid reason.
2. There was a change in how a noninterview outcome was set in the instrument that allowed some 2010 cases to be marked as noninterviews that would have been recorded as partial interviews in 2000. In 2000, the instrument only checked if certain questions early in the interview were asked to set a partial code, while in 2010 all questions in the proper paths had to have some sort of reply for the case to be considered a sufficient partial interview.

3. In 2000, there was a nonresponse conversion phase. This set a time aside at the end of production interviewing to focus just on nonresponse conversion. This phase was dropped for 2010 because managers reported that they did not wait for that part of the operation in 2000 to do conversions and would do the same for 2010. Therefore, it was determined to be not needed, but it may be that this assessment was not correct.
4. In 2000, there was a telephone phase in addition to the personal visit phase. This allowed an earlier start to the 2000 PI operation and may have gotten more completed interviews. Based on observations from interviewers and Headquarters' observers in 2010, it seems respondents were open to telephone interviews.
5. The timing of the 2000 operation was earlier in the census cycle than for 2010. This meant that Census advertising was still taking place in 2000 when the PI operation started. The PI cycle was delayed for 2010 to avoid overlap, and possible contamination, with ongoing Census operations.

We also expect that the extensive 2010 Census advertisement campaign which concentrated on the "Just 10 Questions" slogan could have had an effect on the noninterview rate for 2010 PI, since we did ask substantially more than ten questions. In addition, there is evidence of an overall decline in response rates for government surveys. (Linse, 2010)

One of the things we want to look at is how to best minimize partial interviews. Sufficient partials do count as complete cases, but at the same time we want to minimize them, so we can decrease the overall amount of missing data. Cases are a sufficient partial if names are missing or they refused to answer or did not know one of the residence questions. Overall, there were 11,399 cases that were sufficient partials. Most of the sufficient partials resulted when the respondent could not answer if a person had another type of alternate address for at least one of the people in the household. This situation accounted for 54.98 percent (6,207) of sufficient partial cases. In addition, 33.84 percent (3,857) of sufficient partial cases could report if a person had another address and did report having at least one alternate address, but could not provide any component of the address (e.g., city).

While sufficient partials are not optimal, they do provide a fairly complete picture of the household for matching. More importantly, we need to try to minimize insufficient partials, where we get a respondent but they chose to end the interview before we can complete the case. We call these cases "break-offs". We only got 1,067 cases that are designated as coded "219: Other Noninterviews" and only a subset would be break-off cases. It appears that 949 of these cases never got past the front section on the instrument; in other words, these cases never got a refusal but never reached a cooperative respondent and when the survey ended they were still trying to find someone to cooperate. Only 118 cases have at least one person listed and hence, are considered a true break-off. Of those, 75 stopped the interview before the end of the roster probes and 96 ended the interview by the time the interviewer reached the Outmover question. This most likely means that interview break-offs are not a large issue for PI because most likely these are reluctant respondents that just continued with the interview a little longer than most nonrespondents. Overall, we need to concentrate on the reluctant respondents that do not refuse, but we cannot get to complete the interview. It appears that once we get most people to begin the interview they do complete it.

5.5 How much effort was necessary by interviewers to get each type of status for a case?

Many field procedures are used to lower the amount of effort needed to get a complete interview and lower the cost. Interviewers were supposed to visit households on different days and at different times of the day in order to best catch someone home. If a case appeared to be vacant or not a housing unit, interviewers were allowed to do an immediate proxy to confirm the status and complete the case. They were also allowed to do proxies when no occupant in the sample address was available (either away for the entire survey period or all incapacitated). We also suggested leaving telephone numbers on the "Notice of Visit" to complete the interview over the telephone or to set up an appointment to visit.

Table 16 reviews the number of attempts needed; grouped by one, two to three, four to six, and over six attempts, for each of the four outcome types. It also provides the minimum, the maximum and average number of attempts per outcome type.

Table 16: Person Interview Number of Interview Attempts Plus Average Attempts by Outcome Type

# of Attempts	Complete		Noninterview		Not a Housing Unit Now		Vacant Now		Overall	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
1	68,427	44.08	870	35.12	3,663	61.67	8,173	35.37	81,133	43.44
2-3	55,761	35.92	472	19.06	1,795	30.22	9,177	39.72	67,205	35.98
4-6	20,873	13.45	447	18.05	377	6.35	3,904	16.90	25,601	13.71
Over 6	10,182	6.56	688	27.78	105	1.77	1,852	8.02	12,827	6.87
Average (Min, Max)	2.56 (1, 47)		4.80 (0, 45)		1.76 (1, 43)		2.83 (1, 37)		2.60 (0,47)	

Source: Person Interview and Reinterview Output

Key points from Table 16:

- All of these procedures to limit the number of attempts needed to complete an interview seemed to have worked since we ended up with 79.42 percent of all cases resolved in three attempts or fewer[11]. Overall, 93.13 percent of all cases were resolved in six or fewer attempts.
- Interestingly, Completes and Vacants took on average the same number of visits. This is most likely because the nature of addresses that are Not a Housing Unit are more readily apparent in the fact that they are usually either a business or visibly not able to house people (under construction, burned out, etc.), while for vacants we probably would need a few visits to confirm. So Vacants match up to the number of attempts needed for Completes as compared with matching up with Not a Housing Units.
- Eighty percent of Completes were completed in three attempts or fewer. This shows that getting an interview is still fairly straight forward for most cases and a majority of the effort is only for the hard to reach cases.

[11] For all references to count attempts in this section, including Tables 16 and 17, please see limitations section 4.B.

- 54.18 percent of noninterviews were done in three or fewer attempts. Most likely, these are refusals.
- Noninterviews, as expected, require more attempts in general (4.80 average), compared to the overall average of 2.60 attempts.

For Completes, the distribution of the number of attempts should differ by the type of respondent. Normally, an interviewer is expected to make at least six visits and wait three weeks until moving on to completing interviews with a knowledgeable proxy. However, proxies were allowed early in cases where no adult occupant would be available for the entire survey duration (including incapacitation). This was an exception that should not happen often. This is controlled in the instrument as well. Therefore, we would expect to see for Completes by proxy interviews with fewer than six attempts to be much higher than those with the sample occupant interviews.

Table 17 shows the distribution.

Table 17: Person Interview Number of Attempts for Completes by Sample and Proxy Respondent

Number of Attempts	Sample Occupant Respondent			Proxy Respondent		
	Count	Percent	Cum. Percent	Count	Percent	Cum. Percent
1*	67,205	45.34	45.34	1,222	17.38	17.38
2-3	54,251	36.60	81.94	1,510	21.47	38.85
4-6	19,470	13.14	95.08	1,403	19.95	58.80
Over 6	7,284	4.91	100.00	2,898	41.21	100.00
Total	148,210	100.00	100.00	7,033	100.00	100.00
Source: Person Interview and Reinterview Output						

* For the error in count attempts, if the count was zero, we are assuming at least one attempt was made. See limitations, Section 4.B.

Table 17 shows that 58.80 percent of proxy interviews were completed in six attempts or fewer, it is much lower than the 95.09 percent of interviews done by the sample occupants in the same number of attempts. It may be that the proxy rules were not followed as closely as we expected, but it could also be said that the counts for cases that ended up using a proxy have a better chance of having the wrong count attempts because they would be the more difficult cases and therefore are more likely to have been reassigned or assigned to multiple people. See Section 5.13 for more on the reasons proxies were conducted and their characteristics.

Another way to review the level of effort required to complete a case is to study the length of time spent in the instrument. Interviewers are taught to enter the instrument for every attempt when they try to contact someone. This not only records the count attempts but measures the amount of time working the case at the unit. BLAISE software records two different times; the time spent in the last attempt and the total time of all attempts[12].

[12] For all references to time in this section, including Tables 18 and 19, see limitations section 4.C.

Table 18 provides these two averages by outcome type and overall.

Table 18: Person Interview Average Length of Time in Instrument by Outcome Type

Type of Case	Average Length of Time for whole case (minutes)	Average Length of time for last attempt (minutes)
Complete	16	12
Noninterview	17	2
Not a Housing Unit Now	6	4
Vacant Now	9	5
Overall	**15**	**11**
Source: Person Interview and Reinterview Output		

As you would expect, a complete interview takes the longest amount of time for the last attempt. We had estimated that a typical interview would take 10 to 20 minutes; and we can see that it does average 12 minutes in the last session (the session most likely to be the full interview). Due to the high number of attempts for a noninterview, the total time for noninterviews averages the highest, even more than completes (17 minutes as compared with 16 minutes).

Table 19 shows the average length of the total time and last attempt for completes for both sample occupant respondents and proxies. Just like count attempts, we would expect that proxies for complete interviews would require more time to complete overall, even though the interview itself would not necessarily take longer. Table 19 shows this to be true.

Table 19: Person Interview Average Length of Time for Completes by Sample Occupant and Proxy Respondent

	Length of Time for Whole Case in minutes	Length of Time for Last Attempt in minutes
Sample Occupant Respondent	15	12
Proxy Respondent	24	10
Source: Person Interview and Reinterview Output		

5.6 How effective was the automation of the Person Interview Operation?

Automation has many advantages but also has some drawbacks. Overall, the automation led to a better instrument with less error than if it had been a paper and pencil questionnaire. Some of the advantages are:

- Data are delivered daily to the RCCs and Headquarters, which allowed for more review of data on a flow basis and allows project managers to catch any potential errors at the beginning of the survey and correct either through instrument updates or through field procedures.
- An automated instrument helps ensure the interview followed the proper flow by eliminating interviewer error in following the skip patterns and allowed the question pattern to be more specific to the case based on previous responses, resulting in fewer questions being asked and lowering respondent burden.

- In the instrument design, certain questions can be made "must fill," limiting the amount of missing data.
- No printing of forms and no data capture is needed; saving time and money in the overall CCM schedule.
- The interviewers reported that they enjoyed using the laptop and found it fairly straight forward to use.
- The daily transmission of data to the field offices allows easier and timelier tracking of workload.
- Respondents seem to react more favorably to data being collected on an automated instrument/laptop than paper data collection.

While there are many positives to full automation, there are also some issues. Due to the timing of a decision that CCM automation would be descoped from the 2010 Census FDCA contract, the amount of time available to develop the PI and PI RI automated instruments and their control systems was very limited. This resulted in a very tight schedule for development and testing of the instrument and operational control system that also delayed development of training and procedures. Still, very few technical problems were experienced. Only 2,513 calls were placed specifically for PI at Headquarters' technical help desk. Listed below are the top problems received at the help desk at Headquarters run by the TMO system and in order of how often they occurred:

Coverage Measurement Operations Control System (CMOCS):

1. **Account Management –** Help was requested with Creating, Modifying and deleting CMOCS user accounts.

2. **Procedural Issues –** Help was requested with functions the user was not properly executing. These were not usually true errors; but; the user was not familiar with functionality.
   - *Example 1:* Cannot delete duplicate case in Supervisory Review.
     Resolution: Either accept or reassign one of the duplicate cases and the other one will disappear.
   - *Example 2:* Case keeps appearing on Supervisory Review after many reviews.
     Resolution: Supervisor needs to mark the case as accepted and it will not appear again.

3. **Software Errors –** Issues with the software – True Errors
   - *Example 1:* CMOCS will not load; user is getting the "PB failed" error.
     Resolution: There was an unsuccessful Oracle Client upgrade, which needed to be resolved. Users needed to log out of Novell and log back in to access the CMOCS application.
   - *Example 2:* User getting a database access error when attempting to access CMOCS.
     Resolution: Oracle path was updated and resolved the problem.

Laptops:

4. **Transmission Issues –**
   - *Example:* Interviewer unable to change transmission mode from dial-up to broadband.
     Resolution: Walked FR through setting their password and getting past the keystroke error they were encountering.

5. **Laptop Peripheral Issues** – Correcting settings on laptop not set as user expected or user not correctly executing tasks.
   - *Example 1:* FR reports that their laptop screen goes black when the power cord is plugged in.
     Resolution: Walked FR through changing laptop settings so this does not happen. Issue resolved.
   - *Example 2:* FR reporting that their touchpad is not working on their laptop.
     Resolution: Walked FR through changing the setting for the touchpad with the F8 function key.

6. **Software and Application (Case Management) Issues**
   - *Example 1:* FR reporting that the laptop froze mid-interview.
     Resolution: FR instructed to reboot laptop, issue resolved.
   - *Example 2:* FR reporting that they could not open cases received after recent transmission.
     Resolution: Reassigned cases to FR and issue was resolved.

7. **McAfee Endpoint Encryption Password Resets –** Only the main help desk could reset McAfee Encryption Passwords.

8. **Laptop Kit Recovery -** Incidents/tickets from RCC tracking their attempts to recover laptop kit equipment from FRs.

9. **Imaging Issues –** A laptop is set up through a process called imaging. If this is not properly done, issues may arise.
   - *Example:* IT specialist is getting the "program is blocked by group policy" error when attempting to reimage a CCM PI laptop.
     Resolution: The imaging files are corrupt; reloaded with new files.

10. **Virus Scan Flags**
    - *Example:* Laptop flagged for Virus Scan.
      Resolution: Laptop was scanned and sanitized.

Many of the issues listed above are not technical issues but user error or lack of knowledge. Another problem with automation is that the training needs to be much more extensive and there are many details that a user must remember in order for everything to work properly. This puts more of a burden on manuals and training, and even then there is only so much information training can cover in the time budgeted and there is only so much a trainee can retain in one training session.

Due to the complexity of the instrument paths, flags and quantity of questions, the data are not necessarily always clean even with an automated instrument. Some examples of problems encountered in the data that needed to be corrected in post processing are listed below.

1. Off-Path Data – If an interviewer goes down one path of the instrument and needs to back up and change a previous response, this sometimes will then take him down a different path. In this situation, the instrument would continue to store the now off-path data since the interviewer did not go through those questions again to clean the data [For example, a respondent said yes to having a college address and started to provide a college address and then changed the answer to "No college address". The interviewer then backs up in the instrument and changes that answer from Yes to No. This now creates an off-path college address because the data say there should not be one.] Since we were aware of this issue, we attempted to correct all instances of off-path data for major paths such as the example just given, but we discovered that there are infrequently used paths we did not account for and that off-path data remained in some of the cases.
2. Miscoded Cases – This relates to problem number 1. If the interviewer went back and forth many times making changes in either the front or the back of the instrument, the result could be conflicting data and the instrument would do the assignment of an outcome code based on the first values it processed. This was not always the correct outcome.
3. Hidden Noninterviews – Some of the problems we found were not technical problems but unanticipated occurrences. One of these situations was when the interviewer would not exit the case when a person was refusing or not knowledgeable, but go through all the questions asking each and taking a refusal or "Don't Know" as the answer. We assume this was an attempt to get any information they could from the respondent. Since the instrument computed that every question had been asked, it marked this case as finished as a Sufficient Partial and closed the case. What we would have actually liked to have happen was for this case to stay open and to re-attempt the interview at a different time. This is one of the problems we discovered early and monitored daily. We identified these cases for the RCCs to reassign to try to complete.

Luckily, these instrument problems affected very few cases with only 23 problems found and no one problem affected more than 100 cases in the final output. For a full review of the data problems, logged, and corrected; see Attachment B.

5.7 How often did Headquarters receive multiple versions of a specific case?

When the interviewer transmitted completed cases to CMOCS in the RCC, CMOCS automatically forwarded them to Headquarters. (This is because Completes will almost always be accepted.) The RCC staff reviewed all other types of cases and could review Completes as well. Because of reassignments and assigning cases to multiple interviewers (blitzing), more than one version of a case could be received in the RCC. As these cases came in on a flow basis, the RCC could choose to send a new version of the case as a final version at any time. This

means that for some cases Headquarters received more than one version of a case. We wanted to limit these occurrences, since we were processing on a flow basis and all versions received would need to be reviewed.

Headquarters received more than one version for 3,261 (1.75 percent) cases. The most versions sent for one case was four, but most of the multi-version cases (92.33 percent, 3,011) sent only in one additional version. Interestingly, 202 (6.19 percent) of those cases ended up sending the same version twice. In general, this shows that the functionality to send multiple versions was needed but was used minimally as expected.

5.8 How many production cases failed the quality check and needed be replaced with Reinterview cases? What are the characteristics of those Reinterview cases?

For a full review of the PI RI and PI RI MaRCs (the system used for QC sampling of the PI and response matching of PI and PI RI data), see the 2010 Census Coverage Measurement: Person Interview Quality Profile Early Closeout Draft[13]. (Hartman, 2012) Below we provide a short review of how RI worked and a review of the RI cases that ended up as the final version of some of the PI cases.

RI sampled 16.46 percent of cases from PI. A case was selected for RI in one of four ways: Random, Supplemental, Outlier or Hard Fail. Random RI is a random sample of cases within each interviewer's workload. Supplemental RI are cases that the RCC can select if they would like to review more work of any interviewer. Outlier RI are cases selected weekly when an interviewer is outside the normal case distribution compared to other interviewers in their crew, by reviewing items like vacancy rates and number of one-person households. Hard Fail RI cases were selected when an interviewer had a case that received a match code of Hard Fail. At that time, all eligible cases that the interviewer had worked were sent out to RI as Hard Fail cases that needed to be checked.

The Quality Control (QC) program for PI consisted of an independent quality check via the PI RI interview to verify that the sample address was contacted in PI, if the interviewer collected the correct housing unit status, if they at least rostered the correct number of people, and compared names for minimal differences. The PI RI instrument collects the status of the unit on PI Interview Day, if the respondent remembers having been contacted, and collects a roster of people if the unit is occupied. That information is compared to the data collected in PI and assigned a match code. This match code indicates whether the PI interview is suspected to be incorrect or not and, if incorrect, whether or not they suspect the interviewer error was intentional. The following is the list of possible match codes assigned in RI matching:

- Pass – PI is correct.
- "Don't Know"/No suspect Falsification – Not sure if PI is correct. Do not suspect any intentional falsification.

[13] The 2010 Census Coverage Measurement Person Interview Quality Profile was originally part of the CPEX assessment series. Due to budget restrictions, it was removed. At the point of removal, analysis was done and the draft had been reviewed by the Executive Steering Committee. For the permanent record, it has been stored as a draft within DSSD.

- "Don't Know"/Suspect Falsification – Not sure if PI is correct. Suspect possible falsification.
- Soft Fail – No falsification suspected but interviewer did something incorrect in PI.
- Hard Fail – Falsification suspected (usually confirmed.)

If a case hard fails, then the interviewer's entire completed eligible workload is selected for RI to check all of their work. Only 19 interviewers hard failed for PI[14] and that resulted in 556 cases being resent as hard fail RI cases to check for more possible falsification by those interviewers. Only 60 of those cases resulted in Hard Fail match codes confirming more falsification. Overall, 107 cases were confirmed Hard Fails (i.e., confirmed as falsified).

If a case gets either a Soft Fail or Hard Fail match code, then the final version of the case used in processing is the RI version. If the cases get either of the "Don't Know" codes and they did a full PI interview during RI then the RI version is used as the final version as well. If they did not do a full PI interview, then due to the uncertainty, we keep the PI case as the final version in the hope that this version will contain more of the case's responses.

There were 816 cases delivered to post-processing as cases that did not Pass[15]. Of those, 789 RI cases were selected to replace the PI interview. A total of 669 (84.79 percent) cases were selected to replace PI because they were a soft fail (the information in the interview was confirmed wrong, but it was due to a mistake by the interviewer and not intentional). Those cases are treated just like a PI case in this analysis, but 227 (28.77 percent) cases were a complete RI, but did not ask any of the PI residence questions because during the RI the respondent said they did talk to someone for PI. Therefore, the instrument computed that a full interview was not necessary. This means they had only a roster and no residence status could be computed. People with no residence status are almost always sent to CCM Person Followup (PFU). While we did not replace too many cases with RI, one-third of the RI cases we needed to use as the final form did not have enough information because they did not collect a full interview and that is very limiting in use of the data. For future application, we should review if there are any other indicators that might indicate when we should do a full PI interview, so that we do not have to revisit these cases in PFU like we had to in 2010.

Table 20 compares why a case was sent to RI by the final RI outcome we received for cases that replaced PI as the final version.

[14] Only 18 Interviewers were officially hard failed in the system. One person was not hard failed in the system because the case they marked as hard fail came after the sampling system was turned off.

[15] This file was delivered daily when a Local Census Office (LCO) closed out of RI. This was to ensure that a case would not be updated after matching began. Some LCOs though did change their match codes, requiring the LCO file to be reproduced. Due to this, the initial count of cases is 817 and the final is 816.

Table 20: Total Reinterview Cases that Replaced the Person Interview Version by Selection Type and the Outcome Type

Selection Type	Overall		Complete		Noninterview		Not a HU		Vacant	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
Random	412	52.22	285	50.44	31	65.96	37	64.91	59	49.17
Outlier	272	34.37	192	33.98	9	19.15	19	33.33	52	43.33
Supplemental	32	4.06	28	4.96	1	2.13	1	1.76	2	1.67
Hard Fail	73	9.25	60	10.62	6	12.76	0	0.00	7	5.83
Total	789	100.0	565	100.00	47	100.00	57	100.00	120	100.00
Source: RI Sample Selection File and PI RI Output										

Most (52.22 percent) of the cases selected as the final version were sampled for RI as random cases. In addition, a high "Vacants" rate or "Not a housing unit" rate in PI are some of the parameters to send cases for an interviewer to RI as an outlier. This could be why Vacant RI cases do not have the same sampling pattern for RI as the other categories and overall.

As part of this analysis, we looked at the characteristics of the RI cases selected as the final version and also compared them to the overall PI cases. As expected the average number of people listed in a RI household is very close to the overall average number of people listed in PI (1.9 as compared with 2.3). RI did have a difference in the number of cases that reported an alternate address, with 23.15 percent of PI cases overall reporting at least one alternate address compared to just 14.87 percent (84) of RI cases. This makes sense since about one-third (244 out of 789) of the RI cases that were selected as PI's final version had only the roster collected.

RI needs to be reviewed to see how to make it a more comprehensive quality check. The final field costs (not including RCC office staff) were $ 5,107,797. This means that it cost $6,474 for each case that ended up being replaced and this is just the field cost. We should be looking at ways to make RI a more complete check of the quality and not looking just for falsification and errors in rostering. The RI process also needs a way to make sure that when a case ends up being the final version of PI, it has all the information an original PI would include. If this does not happen, then it needs to be scaled back, since there is not much evidence for falsification or errors that require version replacement and our operation is too short to provide very much feedback based on RI by the time review is done.

5.9 How effective was Person Interview at rostering people?

The goal of CCM PI is to list all people that live or stay at the sample address, both at the time of interview and on Census Day. In the 2010 planning, the CCM design team determined that instead of collecting people who lived at the unit on Census Day and then collecting inmovers; respondents would be more accurate listing people who live and stay at the address at the time of the PI interview and then adding to the roster additional people that have moved out of the address since Census Day. This approach was tested in the 2006 Census Test (Hunter and Nichols, 2004). We call these people who have left since Census Day, outmovers. We can collect data on individual outmovers or, in some cases; the entire household living at the sample address can switch between Interview Day and Census Day. When this happens, we call these

people - Whole Household of Outmovers (WHO). The roster and information on each of these types of people were collected in a different area within the instrument.

Table 21 shows where in the instrument each of these groups of people were collected, and the type of person they were expected to be (Current Resident (Nonmovers), Inmover into Sample Address, Outmover from Sample Address). Most people are collected on the main roster.

Table 21: Person Interview Distribution of People Collected by Instrument Section

Type of Roster	Total Number of People	Percent of People
Main Roster (Current Residents (Nonmovers) and Inmovers)	407,329	96.24
Outmovers Roster	9,168	2.16
Whole Household of Outmovers Roster	6,745	1.59
Total	423,242	100.0
Source: Person Interview and Reinterview Output		

Most PI people (96.24 percent) were collected in the main roster, while the remaining 3.75 percent accounted for rosters collecting outmovers.

Within the main and the outmover rosters, there are different ways to collect people. CCM PI attempted to include anyone who could possibly be counted at the sample address on Census Day, including those that may not consider the sample address their main residence. So, we not only asked about people who lived and stayed at the sample address, but we also probed for the types of people that may stay at the address that are typically missed or may not have been thought of as part of the household. We also collected a roster in situations where the sample address may be transient, based on the type of unit, such a houseboat, recreational vehicle, or a long-term seasonal address, to ensure we collect any people staying there that had no other place to live. Listed below are the main question wording for each roster and probe in the instrument. Table 22 shows the distribution of where each person was collected.

ROSTER 1(Transient): *What are the names of the people staying here and have no other place where they usually live?*
ROSTER 2(Main): *We'll start by making a list of everyone who lives or stays here now. Let's start with you. (Anyone else?)*
PROBE 1: *Is there anyone who has another place to live but stays here often?*
PROBE 2: *Is there anyone who is staying here until they find a place to live?*
PROBE 3: *Are there any babies, foster children, or other children who stay here that you didn't mention yet?*
PROBE 4: *Have I missed any relatives or unrelated people who live or stay here?*
ROSTER REVIEW: *I am going to show you the list of people I recorded. Is everything spelled correctly? Is the list complete?*

OUTMOVER: *Was there anyone else living or staying here during March or April who is no longer living here?*
OUTMOVER REVIEW: *I am going to show you the list of people who moved out or passed away. Have I spelled everything correctly?*
WHOLE HOUSEHOLD of OUTMOVERS ROSTER: *What are the names of the people who lived here on April 1, 2010? (Anyone else?)*

Table 22: Person Interview Distribution of People by Roster/Probe Where Added.

Roster/Probe Where People were Added	Number of People Added	Percent of People Added
Roster 1: Semi-Transient People Roster	240	0.06
Roster 2: People living here now (Main Roster)	396,346	93.65
Probe 1: Stays often	7,909	1.87
Probe 2: No other place to live	1,058	0.25
Probe 3: Other children	1,010	0.24
Probe 4: Other relatives, people, etc.	489	0.12
Roster Review	277	0.07
Outmover Roster	9,153	2.16
Outmover Review	15	0.00
Whole Household of Outmovers Roster	6,745	1.59
Total	423,242	100.00
Source: Person Interview and Reinterview Output		

As expected, most people (93.65 percent) are added in the main roster (Roster 2), but a total of 2.47 percent of the people were added through the four probes and 3.76 percent of people listed were collected by the two outmover rosters. For more on the people collected in the probes, see Section 5.17.4.
Table 23 shows the average number of people listed by outcome categories. As a review of the outcome categories, all completes are occupied housing units on Interview Day and the household size would be those occupants plus any outmovers they listed. For units classified as Not a Housing Unit and Vacant, the average household size is computed based on the roster of people reported that were living at the unit on Census Day. For Noninterviews, the average household size would be expected to be incomplete, but could have a roster if the noninterview is an insufficient partial with a break off after the roster is started.

Table 23: Person Interview Average Household Sizes by Outcome (Plus Minimum and Maximum)

Outcome	Number of Households	Number of People	Average Household Size (Min, Max)
Completes†	155,243	417,706	2.69 (0, 49)
Noninterviews	2,477	381	0.17 (0,10)
Not a Housing Unit	5,940	121	0.02 (0,9)
Vacant	23,106	5,034	0.22 (0,8)
Overall	**186,766**	**423,242**	**2.27 (0,49)**

Source: Person Interview and Reinterview Output

† For PI RI, a case can be complete if a full PI was not required and they at least provided a population count. In two instances, the RI final version had a pop count of zero.

Since, for units classified as Noninterview, Vacant and Not a Housing Unit, we usually have no roster, the number of units with no roster brings the average household size down. The average for Not a Housing Unit is even smaller because that unit status is most likely to have been the same on Census Day, while Vacants have a better chance of having been occupied on Census Day.

We also reviewed the average number of people by the type of respondent for cases that are not a Noninterview[16]. Proxies averaged 0.49 people per case, while Sample Address respondents averaged 2.73 people. This was expected since proxies are less likely to know or provide complete information, are used on Vacant and Not a Housing Unit cases, and may list only the head of household for outmover situations or because they may not know all the household residents. If you look only at Complete Interviews, then the comparison is much closer with proxies averaging 1.76 people listed compared to occupant respondents at 2.73 people.

One of the other things we wanted to consider was if households that have at least one unrelated person were larger or smaller than those where the people in the household are all related. In 2000, on average, unrelated households were smaller than family households, but with the economic situation in the United States in 2010, we suspected that we would get a larger household size for those with unrelated people, since there would be more families sharing a housing unit. Excluding unknown relation households, related households averaged 3.22 people while nonrelated households averaged 1.76 persons. So our hypothesis was not proved to be true, however this is more a factor of considering all one person households as nonrelated. If you ignore the one person households, then the nonrelated households do average slightly more than related households; 3.32 and 3.22 respectively.

Table 24 shows the distribution of the household size in households where a roster was collected (HU either occupied on Interview Day or Census Day) and not a Noninterview. As expected, a majority of the households are in the same grouping as the average number of people, but it is interesting to note that when we separate the two types of rosters the interview collects, the distribution is different with the rosters being created for an entire household of outmovers being

[16] Noninterviews were excluded since it is unknown if the roster was finished before the break-off and respondent information may have not been collected.

just a one-person household. This may be a sign that they are underreporting and only reporting the one person they may have known, instead of all residents. It could also be that single person households are more mobile.

Table 24: Distribution of Person Interview Household Sizes Overall and by Roster Type*

Household Size	All Households		Main and Individual Outmover Rosters		Whole Household of Outmovers Roster	
	Count	Percent	Count	Percent	Count	Percent
1	40,257	25.51	39,381	25.37	1,569	46.53
2-3	75,439	47.81	74,435	47.95	1,345	39.89
4-6	37,990	24.07	37,405	24.09	438	12.99
7-10	3,809	2.41	3,724	2.40	19	0.56
Over 10	306	0.19	297	0.19	1	0.03
Total	157,801	100.00	155,242	100.0	3,372	100.0
Source: Person Interview and Reinterview Output						

*Excluding Noninterview, Vacants and Not a Housing unit where no people were collected.

In order for PI to be most effective both during the interview and when using data in Person Matching, a full name is collected. However, the instrument allows for people to refuse names or to add a person without providing the name. We ask that the respondent give at least a description such as Father or Mother if they do not know or want to give the name.

Table 25 shows the results by the type of names given.

Table 25: Person Interview: Completeness of Name

Completeness of Name	**Number of People**	**Percent of People**
Full Name (at least two characters in both first and last name)	413,455	97.69
Last Name only (at least two characters in last name – should have description*)	1,476	0.35
First Name only (at least two characters in first name)	2,909	0.69
No name – Description only	5,317	1.26
No Name, No Description*	85	0.02
Total	423,242	100.00
Source: Person Interview and Reinterview Output		

*A description of the person was supposed to be filled if first name was not provided, but through user error, the instrument did not collect the description. In addition to the 85 blank names, 16 people did not have a description when only last name was filled for those people.

Overall, 98.72 percent (417,840) of the people had some part of the name reported and could be used in matching. Of those not providing names, 98.43 percent (5,317) at least provided a description that could assist in the Person Matching process.

While the focus of PI is not a distribution by demographic characteristics, this information is needed to best produce proper estimates of coverage and PI is trying to collect data on a representative population of the country. The P-Sample distributions are not comparable to the 2010 Census distributions unless weighted because CCM purposely over samples in certain areas to make sure they have enough data for reliable estimates on certain populations. Table 26 through Table 30 show the distributions of all the demographic characteristics for the people collected in the PI, overall and split between P-Sample and E-Sample. For all the tables, missing includes those people in cases that broke off the interview before reaching that question and those people that answered "Refused" or "Don't Know". Note that the count of E-Sample cases in PI is small, because these are those cases that the Census listed but CCM did not.

Table 26: Person Interview: Sex by Sample

Sex	P Sample (CCM)		E Sample (census-only cases)		Overall	
	Count	Percent	Count	Percent	Count	Percent
Male	197,426	48.22	6,557	47.58	203,983	48.20
Female	206,915	50.53	7,021	50.95	213,936	50.55
Missing	5,121	1.25	202	1.47	5,323	1.26
Total	409,462	100.00	13,780	100.00	423,242	100.00
Source: Person Interview and Reinterview Output and Sample Selection						

Table 27: Person Interview: Age by Sample

Age	P Sample (CCM)		E Sample (census-only cases)		Overall	
	Count	Percent	Count	Percent	Count	Percent
Under 5 years	28724	7.02	958	6.95	29,682	7.01
5 to 9 years	26748	6.53	846	6.14	27,594	6.52
10 to 14 years	26866	6.56	810	5.88	27,676	6.54
15 to 19 years	28236	6.90	886	6.43	29,122	6.88
20 to 24 years	28549	6.97	1,021	7.41	29,570	6.99
25 to 29 years	27105	6.62	906	6.57	28,011	6.62
30 to 34 years	24703	6.03	805	5.84	25,508	6.03
35 to 39 years	24215	5.91	785	5.70	25,000	5.91
40 to 44 years	25203	6.16	829	6.02	26,032	6.15
45 to 49 years	26696	6.52	873	6.34	27,569	6.51
50 to 54 years	26279	6.42	840	6.10	27,119	6.41
55 to 59 years	23125	5.65	709	5.15	23,834	5.63
60 to 64 years	19758	4.83	648	4.70	20,406	4.82
65+ years	45886	11.21	1,823	13.23	47,709	11.27
Missing	27369	6.68	1,041	7.55	28,410	6.71
Total	409,462	100.01*	13,780	100.01*	423,242	100.00
Source: Person Interview and Reinterview Output and Sample Selection File						

*Due to rounding of decimal these columns round above 100 percent.

Table 28: Person Interview: Relationship Status by Sample

Relationship Status	P Sample (CCM)		E Sample (census-only cases)		Overall	
	Count	Percent	Count	Percent	Count	Percent
Householder	153,035	37.37	5,803	42.11	158,838	37.53
Husband or Wife of Householder	70,986	17.34	2,098	15.22	73,084	17.27
Biological Son or Daughter of Householder	113,519	27.72	3,531	25.62	117,050	27.66
Adopted Son or Daughter of Householder	2,479	0.61	50	0.36	2,529	0.60
Stepson or Stepdaughter of Householder	5,611	1.37	187	1.36	5,798	1.37
Brother or Sister of Householder	5,592	1.37	164	1.19	5,756	1.36
Father or Mother of Householder	5,388	1.32	148	1.07	5,536	1.31
Grandchild of Householder	11,573	2.83	376	2.73	11,949	2.82
Parent-in-law of Householder	1,247	0.30	20	0.15	1,267	0.30
Son-in-law or Daughter-in-law of Householder	1,944	0.47	68	0.49	2,012	0.48
Other Relative	7,530	1.84	267	1.94	7,797	1.84
Roomer or Boarder	1,583	0.39	63	0.46	1,646	0.39
Housemate or Roommate	6,808	1.66	245	1.78	7,053	1.67
Unmarried Partner	10,100	2.47	350	2.54	10,450	2.47
Other Nonrelative	7,146	1.75	242	1.76	7,388	1.75
Missing	4,921	1.20	168	1.22	5,089	1.20
Total	409,462	100.01*	13,780	100.00	423,242	100.00

Source: Person Interview and Reinterview Output and Sample Selection File

*Due to rounding of decimal these columns round above 100 percent.

For Table 29, the bold categories are the main Hispanic origin categories and the counts in italics are those for subcategories within that Hispanic only.

Table 29: Person Interview: Hispanic Origin by Sample

Hispanic Origin	P Sample (CCM)		E Sample (census only cases)		Overall	
	Count	Percent	Count	Percent	Count	Percent
Not Hispanic or Latino only	**319,494**	**78.03**	**10,137**	**73.56**	**329,631**	**77.88**
Hispanic only	**81,162**	**19.82**	**3,324**	**24.12**	**84,486**	**19.96**
Mexican only	*40,225*	*9.82*	*1,396*	*10.13*	*41,621*	*9.83*
Puerto Rican only	*21,508*	*5.25*	*1,210*	*8.78*	*22,718*	*5.37*
Cuban only	*1,845*	*0.45*	*99*	*0.72*	*1,944*	*0.46*
Another Hispanic only	*16,139*	*3.94*	*563*	*4.08*	*16,702*	*3.95*
Multiple checkboxes	*737*	*0.18*	*34*	*0.25*	*771*	*0.18*
Missing Specific Origin	*708*	*0.17*	*22*	*0.16*	*730*	*0.17*
Missing	**8,806**	**2.15**	**319**	**2.32**	**9,125**	**2.15**
Total	**409,462**	**100.00**	**13,780**	**100.00**	**423,242**	**100.0**
Source: Person Interview and Reinterview Output and Sample Selection File						

For Table 30, the bold categories are the main race categories and the counts in italics are those for subcategories within that race. The subcategories can add up to more than the main category since the subcategories could have been picked if they were Multiple Races.

Table 30: Person Interview: Race by Sample

Race	P Sample (CCM)		E Sample (census-only cases)		Overall	
	Count	Percent	Count	Percent	Count	Percent
White alone	**271,515**	**66.31**	**8,276**	**60.06**	**279,791**	**66.11**
Black or African American Alone	**42,736**	**10.44**	**1,715**	**12.45**	**44,451**	**10.50**
American Indian and Alaska Native Alone*	**17,663**	**4.31**	**1,130**	**8.20**	**18,793**	**4.44**
Tribal Name Write In Filled	*19,656*	*4.80*	*1,163*	*8.44*	*20,819*	*4.92*
Tribal Name Write In Missing	*1,268*	*0.31*	*42*	*0.30*	*1,310*	*0.31*
Asian Alone*	**21,001**	**5.13**	**614**	**4.46**	**21,615**	**5.11**
Asian Indian alone	*2,587*	*0.63*	*62*	*0.45*	*2,649*	*0.63*
Chinese alone	*4,455*	*1.09*	*145*	*1.05*	*4,600*	*1.09*
Filipino alone	*6,257*	*1.53*	*187*	*1.36*	*6,444*	*1.52*
Japanese alone	*2,817*	*0.69*	*119*	*0.86*	*2,936*	*0.69*
Korean alone	*1,840*	*0.45*	*40*	*0.29*	*1,880*	*0.44*
Vietnamese alone	*2,063*	*0.50*	*34*	*0.25*	*2,097*	*0.50*
Other Asian alone	*2,638*	*0.64*	*74*	*0.54*	*2,712*	*0.64*
Multiple Asian	*1,014*	*0.25*	*35*	*0.25*	*1,049*	*0.25*
Missing Asian	*200*	*0.05*	*11*	*0.08*	*211*	*0.05*
Native Hawaiian or Other Pacific Islander Alone*	**2,456**	**0.60**	**149**	**1.08**	**2,605**	**0.62**
Native Hawaiian checkbox alone	*2,377*	*0.58*	*82*	*0.60*	*2,459*	*0.58*
Guamanian or Chamorro checkbox alone	*140*	*0.03*	*6*	*0.04*	*146*	*0.03*
Samoan checkbox alone	*422*	*0.10*	*52*	*0.38*	*474*	*0.11*
Other Pacific Islander Checkbox alone	*983*	*0.24*	*63*	*0.46*	*1,046*	*0.25*
Multiple NHPI boxes	*128*	*0.03*	*3*	*0.02*	*131*	*0.03*
Missing NHPI	*48*	*0.01*	*0*	*0.00*	*48*	*0.01*
Some Other Race checkbox Alone	**34,361**	**8.39**	**1,225**	**8.89**	**35,586**	**8.41**
Multiple Races	**9,942**	**2.43**	**317**	**2.30**	**10,259**	**2.42**
Missing	**9,788**	**2.39**	**354**	**2.57**	**10,142**	**2.40**
Total	**409462**	**100.00**	**13,780**	**100.00**	**423,242**	**100.00**

Source: Person Interview and Reinterview Output and Sample Selection File

* Main Race Totals are less than subcategories since detail is collected for multiple races as well.

In general, P-Sample and census-only cases have the similar distributions for all the demographic characteristics with the exception of race. E-Sample people have higher percentage of minorities

especially for American Indian and Alaska Native Only. This is not so much of a change in PI but more a product that there are going to be a higher percentage of housing units that do not match during the matching operation on American Indian Reservations due to the lack of city-style addresses and hence will have more census-only cases overall, since non-city style addresses and physical descriptions are harder to match.

The main thing to note for demographic information is that PI had a very low missing data rate with age reporting having the highest at 6.71 percent missing and all others under 2.40 percent.

5.10 How effective was Person Interview at collecting alternate addresses?

In order to determine if a person is an erroneous enumeration or duplicate, we need to collect alternate addresses where people could have been counted on Census Day. Similar to collecting a roster, the interview asks about different types of alternate addresses to make sure we have all the possible addresses where each person could have been counted. Listed below are the different types of alternate address questions we ask and their goal. They are in the same order as in the instrument. See Attachment A for the exact wording.

- Roster Probe Address – When a person is added through a roster probe after the initial roster, this indicates they most likely have another address and we ask for that other address at the time of rostering.
- Inmover Questions – Asks if each person lived somewhere else in March or April. The main goal is to collect the Census Day address for people who have moved into the sample address, but it is worded so that people can also include an address that they cycle to and from.
- Outmover Questions – This collects people who moved out from the sample address around or since Census Day and, when a respondent is knowledgeable, we also collect the information on the address where they moved to and when.

The following are questions relating to shared residency. These addresses were found to be the main type of addresses that caused either missed omissions or duplicates.

- College Address – Parents tend to list children at their home even if they spend most of the time living at college. Determining if people stay at a college address helps identify possible duplication and is also necessary to determine if they are in a GQ.
- Living with Relative (Shared Custody) – This question is mainly looking for people who have children in shared custody but can also include other living situations where a person moves between relative's addresses.
- Military Address – Military personnel may have addresses where they work or train for long periods of time, or could be deployed out of the country for over a year.
- Job Address – This is looking for addresses for people who live or stay somewhere else for work.
- Seasonal Address – This is to collect addresses of any vacation or second homes of the rostered people.
- Other Address – A general question to see if they have any other alternate addresses that may not fit in other categories listed.
- Group Quarters Question – This question asks specifically if they have even spent any time in a GQ and if they were there on April 1, 2010.

The following questions are a review of addresses already collected and additional questions about the timing of where they were at a certain point. For completeness, they can add an address at this point if they did not think of it earlier.

- Verify/Review – In this section, we confirm that the people in the household either have no other address or that we have the correct alternate addresses connected to the correct people.
- Census Day – In this section, we try to determine where each person was living or staying most of the time around April 1, 2010.
- Interview Day – In this section, we try to determine where each person was living or staying most of the time on Interview Day.
- Whole Household Outmover – This collects the address of where an entire household moved to if the sample address contains all inmovers, is currently vacant, or is currently not a housing unit.

Overall, 60,590 alternate addresses were collected. 23.15 percent (43,253) of cases had at least one alternate address and 18.80 percent (79,590) of all people collected reported an alternate address. For cases that had an alternate address reported, the average number of alternate addresses within a housing unit was 1.40 addresses with the maximum number of unique addresses collected in one case being 18. For persons with an alternate address, the average was 1.10 addresses per person with the maximum of five unique alternate addresses per person.

The same address can be reported multiple times within the instrument. The instrument adds an address to a pick list once it is first reported, so that the interviewer does not need to retype all the information if reported for someone else or in another section. An address may be reported more than once because more than one person stayed at it or it may be reported multiple times if the address fits into more than one category (e.g., a college student's dorm can be reported when asking about college and again when asking about GQ). Reviewing not only when an address is first reported, but also looking at the overall number of times any address is reported in a category lets us see how the categories are being used by the respondent. Table 31 shows the percentages of where the addresses were first reported and the overall reporting of addresses by section. For example, 6,404 of the unique addresses were first added in the Roster Probe Section, but a roster probe address was reported 7,483 times. Overall, there were a total of 141,939 reports of addresses, but these contained only 60,590 unique addresses. Therefore, it could be said that each address was reported 2.34 times on average for each case.

Table 31: Where Addresses are First Reported in the Person Interview Instrument

Alternate Address Collection (in order of instrument)	Unique Addresses by Section First Reported		Total Number of Addresses Reported * (including re-reporting)	
	Count	Percent	Count	Percent
Roster Probe Address	6,404	10.57	7,483	5.27
Inmover Address	24,217	39.97	43,673	30.77
Outmover Address	4,284	7.07	5,478	3.86
College Address	1,852	3.06	7,428	5.23
Relative's Address (Shared Custody)	3,713	6.13	11,009	7.76
Military Address	177	0.29	300	0.21
Job Address	935	1.54	1,857	1.31
Seasonal Address	6,262	10.34	11,341	7.99
Other Address	4,114	6.79	12,298	8.66
Group Quarter's Address	1,008	1.66	1,683	1.19
Verify/Review	2,537	4.19	2,552	1.80
Census Day Location	324	0.53	18,422	12.98
Interview Day Location	591	0.98	11,919	8.40
Whole Household Outmover	4,172	6.89	6,496	4.58
Overall	**60,590**	**100.00**	**141,939**	**100.00**

Source: Person Interview Final Version Output

*This total includes address reporting across people and sections.

As expected, percentage reporting changed the most between first reporting and overall reporting for Census Day and Interview Day addresses since these questions are intended to identify at which of the already mentioned addresses they were at the time, instead of collecting more alternate addresses. For these questions, it would make sense that the address was reported earlier in one of the other sections and is now being referenced. If the interview had worked as intended, addresses should not be reported for the first time at this point.

As you can see, the number of addresses in a section for overall reporting is usually 1.7 times the number for first reporting. This is because the address could be not only reported for the same person more than once, but could also be reported more than once for different people within the case. Overall, 33.05 percent (20,027) of addresses were reported more than once for the same person, but 26.03 percent (15,534) of addresses[17] were shared between people as well. The average number of people connected to an address is 1.45 with one address actually being connected to 30 people.

The Inmover section is where the respondent reported alternate addresses for the first time most often (39.97 percent). This makes sense, since for most interviews this is the first time we ask about another address. While it is not a problem that they report any alternate addresses at this

[17] For this number, the count of unique address does not include Census Day and Interview Day addresses since they are not questions for reporting addresses but for reporting a location at a time.

point, the original goal of this question was to collect Census Day addresses for those people who moved in. In the end, 27.51 percent (12,014) of addresses reported in this section were not connected to a move.

While it is good to collect an address and to know whether a person should be counted at the sample address or at the alternate address, for the address to be most successful for matching and ultimately estimation, we need to be able to geocode that address. Geocoding is the first process that occurs in the Person Matching Operation after the PI data are processed. The geocoding is done in two steps. First, an alternate address is sent through computer geocoding where all complete addresses have a good chance of being matched to the existing Census address list. Then, all other addresses are clerically geocoded. The optimal goal is to match to a census unit (mapspot), but if that cannot be done, to at least be able to identify the census block(s) that the address is most likely within. For more on how successful geocoding was see "CCM Operational Assessment for Person Matching and Person Followup" (Sanchez et al.) listed in the related evaluations.

Since every bit of information helps for geocoding, we had two ways we tried to get more information on the alternate address; probing and collecting markers. When first reporting the address, if the respondents said they did not know the address, we probed to see if they knew any parts at all. The probe was used a total of 8,640 times and 40.75 percent (3,521) of the time it led to at least one piece of the address being collected, 5.81 percent of all unique addresses reported. When looking at the type of data collected when probing was used, usually only a city and state was given.

Also after collecting the address, we asked for markers to help us try to locate and geocode it. These markers consisted of three things: landmarks, cross streets, and neighbors' names. Of course, what we wanted the most was a complete city-style address. (See Section 5.15 for Puerto Rico address review).

Table 32 shows some of the distribution of the pieces of addresses collected excluding cases in Puerto Rico. We do not need a complete address when the address is outside of the United States or in Puerto Rico for Stateside addresses (Row 1) because those are excluded from the CCM processing. The information on Rows 2 and 3 should let us match to a specific address. Row 4 should allow us to locate the census block(s) in most cases. Often, we would have only the city and state though (Row 5). The other two rows show the other combinations and when we had absolutely no information.

Table 32: Completeness of Alternate Addresses Collected in Person Interview*

Completeness of Alternate Addresses		Count	Percent
1	Outside the Country or In Puerto Rico	2,569	4.40
2	House Number – Street Name – City – State – Zip Code	27,046	46.34
3	House Number – Street Name – (City, State or Zip Code in some combination)	3,070	5.26
4	Street Name – State – (City or Zip Code)	5,914	10.13
5	State – (City or Zip Code)	12,204	20.91
6	All other combinations	3,041	5.21
7	No Parts	4,518	7.74
	Total	**58,362***	**100.00**
Source: Person Interview Final Version Output			

*excluding Puerto Rico (See section 5.15)

From Row 2, we can see that 46.34 percent of the Stateside alternate addresses collected were complete. We would consider Rows 2 through 4 probably geocodable – 61.73 percent of addresses. The markers can help us geocode for incomplete addresses. Table 33 shows the number of times the markers were provided overall.

Table 33: Completeness of Markers Collected in Person Interview*

Marker	**Filled with a Valid Answer**		**Marked Don't Know/Refused/ Left Blank**	
	Count	**Percent**	**Count**	**Percent**
Landmarks	28,170	46.49	32,420	53.51
Cross Streets	27,893	46.04	32,697	53.96
Neighbors	8,409	13.88	52,181	86.12
Source: Person Interview Final Version Output				

* Since Markers are same for both Stateside and Puerto Rico, Puerto Rico is **not** excluded.

When we look to see how often we got neighbors when we did not get a complete address (i.e., Rows 4, 5, and 6 in Table 31), then the percent providing a valid answer drops even more to just 6.68 percent (1,618). Considering how little people knew or were willing to provide the neighbor's names near the sample address, we may want to consider how often matching did use this information and consider dropping the question in the future. Through debriefings with the interviewers, we discovered that these three questions were actually quite sensitive for respondents and they did not want to answer. Since we saw in matching that cross streets and landmarks were highly useful, we may want to consider rewording to try to make these questions less sensitive.

In addition to knowing the address and where it may be located, we needed to know if the unit was a GQ or not. GQs are handled differently in determining where a person should be counted in the Census, so it is important to know if the alternate address they listed is a GQ. For some types of addresses, it was assumed that the address they added such as a relative's address or a second home was a housing unit, but for most of the sections we asked questions to see what type of place the address was, i.e., if it was a GQ. Table 34 shows for each questionnaire section the distribution of responses when the Type of Unit question was asked. In some sections, we ask type only if it is the first time collecting the address, while in other sections, we ask for all

alternate address collection even if the address was reported before. Table 34 represents every time the type of unit questions were asked.

Table 34: Person Interview Type of Unit for Alternate Addresses by Collection Section

Collection Section	Housing Unit		Group Quarters		Don't Know	
	Count	Percent	Count	Percent	Count	Percent
Inmover	19,555	80.75	2,201	9.09	2,461	10.16
Outmover	3,740	87.30	380	8.87	164	3.83
College	4,259	57.34	2,692	36.24	477	6.42
Military	110	36.66	170	56.67	20	6.67
GQ	384	22.82	1268	75.34	31	1.84
Live Now	383	64.81	68	11.51	140	23.69
Household Outmover	2,235	53.57	137	3.28	1,800	43.14
Total	30,666	71.86	6,916	16.21	5,093	11.93
Source: Person Interview Final Version Output						

When the type of unit was collected, 16.21 percent of addresses were GQs. The percentage of addresses for which respondents could not report the type of unit is surprisingly high (11.93 percent). However, if you ignore the Household Outmover section, which is more of a representation of people not knowing where the household moved to at all, then the overall missing rate drops to 8.06 percent (3,293 of 40,875). Either way, this may be a sign that respondents are not understanding what a GQ is even with the handout that has a list of typical GQs.

In general, PI did a very good job of collecting alternate addresses. Overall, 23.15 percent (43,253) of cases had at least one alternate address and 18.80 percent (79,590) of all people collected reported an alternate address.

5.11 How did Person Interview do at collecting the timing of when people were living or staying at the Sample Address or Alternate Address?

In the end, we hope to get sufficient information from PI to be able to determine where people should be counted on Census Day. The more complete the PI information, then the smaller the PFU workload and cases will have less missing data for the coverage estimates. In the previous section, we looked at how well we could identify other possible addresses. Now, we need to review how often the respondent in PI can report if the person should be counted at the sample address or at some other address. We know where a person should be counted based on time of staying at the address reported. Therefore, there are four main scenarios or rules that determine where a person is counted:

1. The person has no alternate addresses reported, so they should be counted at the sample address.
2. The person has a GQ as an alternate address and stayed there on Census Day.
3. The person has one alternate address and they reported they moved around or since Census Day. Use the move date to determine where the person should be counted.

4. The person has one alternate address they go back and forth from (cyclers) or the person has more than one alternate address. Determine where they spent most of the time around Census Day.

For scenario one, since there is nowhere else the person could be counted, residence and timing is very clear. A total of 343,652 (81.20 percent) of people were in this category.

For scenario two, anyone staying at a GQ on Census Day is counted at the GQ, regardless of how long they stayed there or if they spent most of time somewhere else. Only 5,762 (1.36 percent) of people reported staying at a GQ around Census Day in 2010. If they did not know if they were in the GQ on Census Day, then we cannot tell where they should be counted, so they remain unresolved.

Table 35 shows the distribution of people with a GQ address by whether they were at the GQ on Census Day.

Table 35: Person Interview Census Day Status of People with Group Quarters' Alternate Addresses

At GQ on Census Day?	Number	Percent
Yes, On Census Day	1,725	29.94
Not On Census Day	1,112	19.30
Don't know if on Census Day	2,925	50.76
Total	5,762	100.0
Source: Person Interview Final Version Output		

Although 29.94 percent said they were at the GQ on Census Day and 19.30 percent were not, we still had 50.76 percent that did not know if they were there on that exact date. The high percentage of "Don't Know" responses come mostly from two sections that collect GQ information, but do not specifically ask if they were there on Census Day. This leads to collecting where they spend most of their time around Census Day, but we cannot confirm that they were there exactly on Census Day. This situation accounts for 2,753 cases where we do not know if they were there on Census Day (94.12 percent of the "Don't Know" responses). This was a flaw in the instrument logic and would need to be corrected in future uses. If they were not at the GQ on Census Day, then they follow either rule 3 or 4.

For scenario 3 on the list, we need to determine the move date for anyone who indicated that they moved. Without an accurate move date, we at least need to know if the person moved before or after April 1, 2010. If the exact date is unclear, we ask the respondent if they know at least this information. Table 36 shows how often the respondent could provide a move date by type of mover.

Table 36: Person Interview Reporting of Move Date by Where Move Date was Collected*

Section of instrument / Answer	Inmover		Outmover		Move Question† (Inmover)	
	Count	Percent	Count	Percent	Count	Percent
Knew Full Date	27,981	88.34	10,730	68.97	3,518	76.38
Knew Enough of Date to determine Before/After April 1. (i.e. Not during April 2010)	2,091	6.60	1,720	11.06	600	13.03
Knew if before or after April 1st based on Question	1,357	4.28	2,672	17.18	236	5.12
Did not know	249	0.79	436	2.81	252	5.47
Total	31,675	100.00	15,558	100.00	4,606	100.00
Source: Person Interview Final Version Output						

* For more complete review of collection of move date, Table 35 contains all reporting of moves even if they also had to have the most time questions asked due to other alternate addresses.
†See Section 5.16.3 for limitation on the Move Question

It makes sense that a respondent would be more likely to know the move date for an inmover that currently lives with the respondent at the sample address (88.34 percent compared to outmover at 68.97 percent). Interestingly even for outmover dates, respondents were able to at least report if the move was before or after April 1, 2010; keeping the truly missing move period to a minimum of 1.81 percent overall.

For scenario 4 on the list above, we need to know where the person (cycler) spent most of the time. If a person could not determine where they were most of the time (because it was almost equal time), we count the person where they stayed on Census Day. Table 37 shows a distribution of how the respondent reported where they lived most of the time when they were asked the question. (i.e., mainly people who have more than one alternate address, have one alternate address but did not indicate a move, or marked the living situation as complex in an earlier situation.) For completeness, Table 37 reports all people asked the cycle series of questions.

Table 37: Person Interview: Where Cyclers were Counted on Census Day

Where answers indicated they should be counted	Count	Percent
Count at sample address on Census Day	23,198	55.99
Count at another address on Census Day	16,889	40.76
Don't know where to count on Census Day	1,345	3.25
Total	41,432	100.00
Source: Person Interview Final Version Output		

Even though over half (55.99 percent) of cyclers were at the sample address on Census Day, it is interesting that of the people that needed to be asked about their multi-address living situation, so many people (40.76 percent) actually would be Census Day residents at an alternate address. We reviewed the characteristics of the people counted at the sample address with those counted at the alternate address and found two categories worth noting. The people collected by the roster probes account for 30.35 percent (5,126) of the people counted at an alternate address while 8.87 percent (2,057) of the people counted at the sample address were added from the roster probes. Also 13.55 percent (2,288) of the people counted at an alternate address are nonrelatives as compared with only 7.92 percent (1,837) of people who should be counted at the sample address.

Table 38 shows the overall results when putting all the people together using either their move date, cycler information, or their indication of where they were on Census Day. While to assign a true residence code we would need to consider where (as in the actual address) they should be counted, this provides a good indication that PI was very successful in collecting the timing when people stayed at reported addresses. We could not obtain the timing needed to determine residence for 1.17 percent of all people we collected.

Table 38: Person Interview Distribution of Where People Should be Counted on Census Day

Where on Census Day?	Number of People	Percent of People
Count at the Sample Address	373,634	88.28
Count at an Alternate Address	44,654	10.55
Don't Know	4,954	1.17
Total	423,242	100.00
Source: Person Interview Final Version Output		

5.12 How effective was Person Interview at collecting information to determine where people should be counted on Census Day?

In Section 5.10, we reviewed collecting alternate addresses. In Section 5.11, we reviewed how to determine where people should be counted on Census Day based on timing. When you put all this information together, we create a residence status code. This code indicates where we think each person rostered should be counted at the end of PI. This is not the final residence status code since all the information will be reviewed in person matching and many people will go through the PFU operation.

Table 39 shows the distribution of how each person was coded at the end of PI. These codes were assigned based on the data collected, but we also marked more people for review than the data may have required. These were reviewed for the sake of completeness and accuracy. The following types of people get coded for review if:

- Any of the data contained a note
- They were a cycler and the timing indicated they should be counted at an alternate address
- Any of the main questions about residence had a "Don't Know" or "Refused" answer
- Any data are not complete for an Inmover or Outmover including address
- Any of the type of housing unit information was missing.

The assignment of the "Review" code was designed to be on the more conservative side to make sure anything that could be erroneous or not complete was reviewed.

Table 39: Person Interview Residence Status Code

Residence Code Values	Number of People	Percent of People
Inmover – Person is a Inmover	3,579	0.85
Nonmover – Person was at Sample Address most of the time around Census Day	330,436	78.07
Outmover- Person was an outmover*	3,723	0.88
Review – Review in Matching is needed	78,023	18.43
Unresolved – Not enough data were collected to resolve this person	4,015	0.95
Out of Scope – The person should be counted in a GQ, out of the country or died before Census Day	3,466	0.82
Total	423,242	100.00

Source: Person Interview Final Post Processing Output

*For the purposes of this review, the two outmover values (O, Z) have been combined. O is for if a person moved to an in-scope address for CCM and Z is for persons who moved to an out of scope address -- a GQ or outside the country.

Post Processing ended up assigning a good residence status code (not Review or Unresolved) to most of the persons reported in PI (80.62 percent), while 18.43 percent of the people were sent to clerical review to assign a residence status code. If the people with an "Unresolved" code had a complete name, they were sent to PFU. Other codes could be sent to PFU if necessary based on what was reviewed and the results of matching. See the "CCM Operational Assessment for Person Matching and Followup" for more on the results of the clerical residence coding and what was sent to PFU.

5.13 What were the missing data rates for residence questions? What other questions had a high missing data rate?

To be sure we have accurately collected all possible alternate addresses; there are certain questions within the PI interview that must be answered to assign a residence code with confidence. These include the main residence questions and the questions involving determining the dates of staying at addresses. Keep in mind that not all these questions are supposed to be asked of all people. Most of them are restricted by age, e.g., you would not ask the college question to someone in their eighties. The rates in Table 40 show how often the data are missing when the question should have been asked. These rates are a combination of those with "Don't Know" and "Refused" answers and exiting the interview before finishing. This table also shows results by occupant and proxy respondents to show the comparison.

Table 40: Person Interview Missing Data Rates for Main Alternate Address and Residence Questions

Main Questions Used in Assigning Residence Code	Occupant Respondent Missing Rate	Proxy Respondent Missing Rate	Overall Missing Rate
Inmover: Somewhere else around April 1	0.32	5.27	0.46
Outmover: Anyone else that was living here on Census Day	0.57	15.02	1.21
College: Away for College	0.80	38.00	2.77
Shared Custody: Away at a relatives	0.51	29.33	1.81
Military: Away for Military	0.33	27.27	1.51
Job: Away for a Job	0.37	31.10	1.66
Seasonal: Rent or Own a seasonal Address	0.59	32.60	2.04
Other: Anywhere you stay often	0.55	31.39	1.95
Group Quarters: Stay in Group Quarter	1.31	34.65	2.27
Verify/Review: Verify your other addresses	0.24	6.70	0.48
Live Now: Where you are staying most of time now	0.83	9.91	1.11
Census Day: Where you are most of the time around Census Day	2.05	19.79	2.77
Census Day HU Status: What was status of sample address on Census Day	36.84	11.40	17.54
Source: Person Interview Final Version Output			

* Counts are different for every cell for both asked and should have asked (numerator/denominator).

As expected the proxy respondents were not able to provide as much detail as the current occupants, but it is concerning to have a missing rate over 30 percent on many of the key questions asked to proxies. On the other hand, it is very good to see that for most key questions the missing rate is less than one percent when talking to a current occupant. The biggest exception is Census Day Housing Unit Status (36.84 percent) missing for occupant respondents. However, this question is asked only to current occupants when it is an entire household of inmovers. In this scenario, it makes sense that this type of household would not know the unit

status before they moved in, but a proxy such as a neighbor would likely know this information. (We do not go to a proxy to get this information if the occupant respondent could provide all other data.)

The overall college missing rate is higher than expected (2.77 percent). This likely is not an indication that the respondent was not sure if the people were away for college but more that the April time reference in the question overlaps with when a typical college student could return from college and maybe they were not sure if the student was at college during that period. They could have also responded "Don't Know" if they already knew that they could not provide the information for the college address.

When reviewing other fields in the instrument, there are not many that have a high missing rate overall besides those we have reviewed earlier in this assessment. Demographic information missing rate is around two percent except for age, which is 6.71 percent (See Table 26 through Table 30). The highest missing rates in the instrument were for the markers to help locate the address (See Table 33).

5.14 What impact does using proxies have on data quality?

In order to get a higher response rate, we allowed a proxy to be used if it appeared that the interview was not going to be completed with a household occupant. For refusals, interviewers were not supposed to use a proxy. Before contacting a proxy, interviewers normally had to wait until six attempts and three weeks passed. There were a few exceptions to this rule, where the interviewer could do the interview with a proxy immediately. The immediate proxy scenarios are:

- The house was either currently vacant or not a housing unit.
- They confirmed that all the occupants would not return until after the end of CCM PI.
- All occupants were incapacitated and could not be interviewed.
- There was a special situation where the CL deemed that the interviewer should move on to attempt the interview with a proxy.
- The Reinterviewer could not get the original PI respondent and needed to select a proxy.

In the end, 19.37 percent[18] (36,175) of the total cases were completed by a proxy respondent and 80.16 percent (28,999) of those were for vacant and not a housing unit where a sample occupant was not even available. Therefore, proxies were only used 4.62 percent (7,176) of the time for complete interviews excluding the vacant and not a housing unit cases. Table 41 shows the distribution of the reasons why a proxy was used.

[18] Proxy rates can be calculated by two different variables. One is the Attempt Type that is indicated at the beginning of the instrument and the second is using Respondent Type collected at the end of the instrument. As explained in Limitation A (Section 4), the interview produced many more complex situations than was expected and the attempt type and respondent type did not always match on final output. The respondent type did not match 0.7 percent of the time. After review, we feel that Attempt Type is correct more often when the two variables do not match and have used Attempt Type in all proxy calculations.

Table 41: Person Interview: Reasons for Completing with a Proxy Respondent

Proxy Reason	Count	Percent
Vacant	23,069	63.77
Not a Housing Unit	5,930	16.39
All Occupants Away until after end of PI	2,615	7.23
All occupants incapacitated	1,340	3.70
Proxy Flag Checked by CL/ Time Limit Met*	3,034	8.39
RI needed a proxy	147	0.41
Unknown (Data missing)	40	0.11
Total	36,175	100.00
Source: Person Interview Final Version Output		

* The flag that gets set when the proper time has passed is stored in the same place as the Crew Leader approval flag, so they cannot be distinguished, but based on debriefings CLs claim they did not need to use it much.

As you can see the two reasons to do a complete interview early with a proxy (all occupants away or incapacitated) were chosen 10.93 percent (7.23 percent plus 3.70 percent) of the time a proxy interview was conducted. This was unexpected and it may be that some of the interviewers may have been using this as a work around on cases they wanted to complete earlier than the three week waiting time allotment.

Some of the following are key differences observed in data when using proxies:

- Proxies averaged 0.49 people listed per case while occupant respondents averaged 2.73 people listed per case excluding cases that are noninterviews[19]. If you also exclude Vacant and Not a Housing Unit cases as well and look only at Complete Interviews, then the comparison is much closer with proxies averaging 1.76 people listed compared to the average for occupant respondents of 2.73. This is expected since proxies may be listing only the householder and because they are most commonly used in move situations where they may not know the people living at the sample address from before, except maybe a last name. The distribution of units with rostered people shown in Table 42 reinforces this finding with 53.10 percent of a proxy interviews report only one household member while the proportion for non-proxy interviews is 23.73 percent.

[19] Noninterviews were excluded since it is unknown if the roster was finished before the breakoff and respondent information may not have been collected.

Table 42: Person Interview: Household Size by Type of Respondent

Household Sizes	Percent of Interviews with Occupant	Percent of Interviews with Proxy
1	23.73	53.10
2-3	48.53	36.64
4-6	25.00	9.73
7-10	2.54	0.52
Over 10	0.21	0.01
Source: Person Interview Final Version Output		

- There were differences in the number of alternate addresses collected when looking at respondent type. Table 43 has a summary of the different ways to measure alternate address collection.

Table 43: Differences in Person Interview Alternate Address Collection by Respondent Type

Measures of Alternate Address Collection	Proxy		Occupant		Overall	
	Count	Percent	Count	Percent	Count	Percent
Cases with Alternate Addresses	3,854	10.65	39,399	26.22	43,237	23.46
People with Alternate Addresses	7,403	42.12	72,187	17.79	79,563	18.82
People with Alternate Addresses for Complete Interviews Only	2,456	19.78	72,160	17.80	74,616	17.86
Source: Person Interview Final Version Output						

Overall, 23.46 of cases had alternate address reported. However, only 10.65 percent of cases completed by proxy have an alternate address while 26.22 percent of cases completed by a sample occupant have an alternate address. This is mainly because there were fewer people reported in proxy cases. However, when looking at the percentage of people that have alternate addresses by respondent type, the percentages are reversed. Overall, 18.82 percent of people have an alternate address, but 42.12 percent of people listed by proxies have an alternate address while only 17.79 percent of people listed by sample occupants have an alternate address. This difference is easily explained since proxies mostly are taken for vacants and not a housing unit cases, where the only people listed are the people who moved out. People listed in these types of cases have been confirmed as movers and would have another address in the interview. If we remove vacant and not a housing unit cases and look only at completes, then the difference drops to 19.78 percent for proxies as compared with 17.80 percent for sample occupants.

- While a proxy may be able to tell us that the person or household did have another residence, 43.03 percent (2,320) of the addresses they reported did not provide any part of an address (compared to only 4.15 percent (2,198) for occupants). Also only 26.44 percent (1,426) of addresses reported by a proxy were probably geocodable (See Table 32 for definition) as compared with 65.33 percent (34,606) for those provided by occupants.

- Another deficiency of proxy data are that proxies do not provide as much of the demographic information for sample address occupants. Table 44 shows the missing data rates for the different demographic characteristics.

Table 44: Person Interview Missing Rate for Demographic Characteristics by Respondent Type

Demographic Characteristics	Missing Rate for Proxy Respondents		Missing Rate for Occupant Respondents		Overall	
	Count	Percent	Count	Percent	Count	Percent
Sex	972	5.53	4,351	1.07	5,323	1.29
Age	9,177	52.20	19,233	4.74	28,410	6.71
Hispanic Origin	2,886	16.40	6,238	1.54	9,125	2.15
Race	2,925	16.64	7,217	1.78	10,142	2.40
Relationship	999	5.68	4,090	1.01	5,089	1.20
Source: Person Interview Final Version Output						

It is surprising that there is such a large missing rate for age (52.20 percent) because there is an instruction on the screen to have the respondent make their best guess as to age if unknown. The other larger difference between occupant and proxy reporting for race may not be as surprising since the interviewers reported that race was a fairly sensitive question and also confusing to those of Hispanic origin. Proxies may be even more unwilling to answer this question. There was also a large difference in Hispanic origin.

In addition, proxies had a missing data rate of 20.01 percent (1,407) for tenure while occupant respondents only had a missing data rate of 1.81 percent (2,681).

- Another clear indicator of how little a proxy knows or is willing to tell us about someone not in their household are the very high nonresponse rates for the main residence questions (See Table 40). Missing rates for questions answered by proxy respondents were around 30 percent while for occupant respondents it was usually under one percent. If any of those questions are unanswered that person's residence status is in question and will almost always go to PFU as long as the roster person has a complete name.

- Another thing we would like to collect from proxies is their information, so we have it to return to for either PI RI or PFU. The missing data rates are much lower for their own information. We have a missing name only 3.36 percent (1,216) of the time. They do not provide their address 11.16 percent (4,037) of the time. However, they are not as good about providing their phone number, probably to avoid being contacted again. They did not provide a phone number 21.66 percent of the time (7,836).

Overall, proxies can indicate that there were people living at the sample address on Census Day, but often they provide such limited information that it leads to a case with substantial missing data. In the end, 18.54 percent (6,447) of the CCM PI people sent to PFU were first listed by a proxy in PI. This probably would be even higher but an additional 20.71 percent of people listed by proxy (3,639 people) did not have a full name to be sent out. We should closely review the costs and benefits of collecting proxy data to determine if the limited data we get through them sufficiently adds to the quality of CCM results to justify in the future all the extra time and resources needed for proxy field data collection and processing.

5.15 How often was the interview conducted in Spanish when Stateside? How often was the English interview used in Puerto Rico?

Quite a bit of development time went into reviewing and properly implementing the PI instrument in both Spanish and English. This was done for both Stateside interviews and Puerto Rico interviews. Due to the amount of resources and time needed, we wanted to review how often the translations were used and consider if it is really required, in particular the Spanish version for the U.S., in future operations.

Table 45: Person Interview: Language Reported by Stateside and Puerto Rico (excluding Noninterviews)

Language	Stateside		Puerto Rico	
	Count	Percent	Count	Percent
English	170,377	96.63	45	0.56
Spanish	4,121	2.34	7,890	98.98
Other	1,011	0.57	2	0.03
Missing/By Observation	809	0.46	34	0.43
Total	176,318	100.00	7,971	100.00
Source: Person Interview Final Version Output				

Table 45 shows that Stateside interviews were completed in Spanish 2.34 percent of the time as reported by the interviewers[20]. The English translation in Puerto Rico was hardly used at all; with the Spanish interview being used in Puerto Rico for 98.98 percent of cases. We reviewed Spanish use by State, RCC, and Hispanic Origin as well. The full tables are in Attachment C, but a summary of the key findings follow:

- Overall, Los Angeles and New York RCCs had the highest percent of interviews done in Spanish (8.18 percent and 5.18 percent respectively). However, due to sampling, some areas have more Hispanic population represented. If you look at the number of Spanish Interviews by the percentage of households that contained a person of Hispanic Origin, then it changes to Atlanta and Los Angeles RCCs with 30.26 percent and 24.76 percent respectively, and New York third with 20.53 percent.

[20] This is based on a flag the interviewer selects at the end of the interview to indicate how the interview was conducted. It is not linked to toggling the instrument to the Spanish translation. So the Spanish translation is no good way to confirm the translation was used, but it is assumed.

- When doing the same measures by state, California (7.79 percent), Nevada (6.30 percent), and Florida (6.11 percent) are the top three states with the highest percentage of Spanish Interviews over total interviews. Thirty-six states had less than one percent of their interviews done in Spanish, of which five had no Spanish interviews.

  When looking at the number of Spanish Interviews compared to households that have a Hispanic individual reported, Iowa (44.90 percent), Florida (36.05 percent) and Nevada (28.61 percent) have the highest rates. Iowa is most likely an interviewer influence due to their small number of households with Hispanic Origin (63) and the interviewer always doing a Spanish interview before even attempting an English interview. In addition, 15 states had less than 5 percent of the cases with a person of Hispanic Origin done in Spanish.

- Looking at households with Hispanic Origin across the nation, 23.15 percent of households that were all Hispanic did a Spanish Interview (3,466 of 14,975), but the percentage drops to 16.65 percent when you look at the number of people who did a Spanish interview where the household contained at least one member that was Hispanic. (3,693 of 22,175)

So we see that Spanish Interviews are needed for interviews with some Hispanic populations, but the actual number of interviews done in Spanish overall is very limited, even including Puerto Rico's 12,011 cases.

One of the other things to consider is that throughout debriefing with Spanish speaking interviewers, they consistently had many complaints about the translation due to the Spanish language variations used throughout the nation, as well as by the different Hispanic populations (i.e., Cuban, Mexican, etc.). It is difficult to agree on a translation alternative that will fully satisfy all possible respondents. However, it appears the translation was used, but it was used more as a guide rather than read verbatim as trained, with interviewers making changes to wording as they thought was needed for the area or dialect. While we stressed to read the translation exactly as worded just like the English interview, most likely this "update as needed" technique was the most useable option and should be reviewed to see if this is a valid technique in the future. This could also make developing the Spanish Translation easier with not needing to get as much agreement between dialects. For Stateside, there was also a suggestion to limit the number of supporting materials in Spanish, since they were used so little as well. On a positive side, interviewers said it was very helpful to be able to show the question wording on the laptop to a Spanish-speaking respondent.

5.16 What were the differences and difficulties of conducting Person Interview in Puerto Rico?

For Puerto Rico, PI required a few changes. The most obvious being that training, reports, letters, and the interview needed to be translated or conducted in Spanish. Interviews were mainly conducted in Spanish, so the Puerto Rico instrument opened to the Spanish screens.

Not only did it need to be in Spanish but in few situations, it needed to be updated to the Spanish preferences that are more common in Puerto Rico. This led to a need for almost double review of translations both at Headquarters and in the Boston RCC[21]. This was also an issue for any general public use documents translated at Headquarters that all operations were to use across all geography and operations. Puerto Rico translators found issues in the translations of some of these materials and they could not be changed.

Besides the translation, the other major changes to accommodate for Puerto Rico were updating the instrument, CMOCS, and training with the differences in the address layout. Puerto Rico does not use the same standard address parts as Stateside United States and actually has more ways of designating an address. The address parts are listed below. The dark bullets are the new Puerto Rico specific fields:

- House Number
- Street Name
- Urbanization/Condominium Name – An Area, Sector or Development Name within a geographical area to make a street address unique within the city.
- Complex Name/ID – Multi-unit structure unique within the city
- Unit Designation
- Area Name 1 – The Area Type and Name for the Address – Usually a Barrio or Sector
- Area Name 2 – The Area Type and Name for the Address that is additional to Area Name 1 – Usually Barriadas, Sectores, Parcelas or Comunidades (a subset of Area Name 1)
- Ramal – Arterial Road
- KM/HM – Kilometer/Hectometer – Identifies street distance marker in the more rural areas.
- City (also known as Municipo)
- Zip

Much as we define House Number Street Name, City, State as a complete locatable address (good city-style in the U.S.), Puerto Rico has combinations of address components that are considered complete. Those combinations are:

- House Number +(Street Name or Ramal) and ZIP Code
  - With Urbanization Name and Unit Designation if they exist
- Condominium Name + Unit ID+ ZIP Code
  - With Complex ID and Street Name if they exist
- Area Name 1+ House Number +Street Name + ZIP Code
  - With Area Name 2 and Unit ID if they exist

All other components are just additional information, like KM/HM, which is used like a physical location indicator. Using these definitions, only 28.59 percent (637) of the alternate addresses we collected in Puerto Rico were complete. Through clerical review

[21] All adaptation was done at Headquarters where the main translation resources were for development. Once the adaption and translations were done, all materials were sent to the Puerto Rico Office in the Boston RCC for review and change as necessary for best use in Puerto Rico. This required an even earlier start to training and material development and put extra burden on the Puerto Rico Office that was already doing production work for the earlier CCM operations.

though, this number may be much higher once the parts were compressed and reported as one variable. See the "CCM Operational Assessment for Person Matching and Followup" for results of geocoding Puerto Rico addresses. This is because when reviewing the addresses it is clear that the interviewers did not always make sure parts were in the proper field since different components were reported in any field where there was room.

When comparing to Stateside collection, there is not a big difference in the number of addresses collected or the number of people collected per household. The one thing to point out as a difference is that Puerto Rico had the highest response and cooperation rate of any RCC.

Puerto Rico did lead to one of the larger problems that needed to be resolved during the PI operation. Spanish characters were supposedly not allowed for collection in our system. In Puerto Rico though, the interviewers are familiar with the numeric short keys on the laptop that insert Spanish characters. Since none of our systems were prepared to receive these data, it created problems in storing data and forwarding the data onto the RI instrument. This problem had to be corrected immediately in order to process cases in RI and in Post Processing[22].

5.17 How did components of the instrument work?

5.17.1 Ability to update the sample address

Some of the addresses collected during Independent Listing were only physical descriptions if the lister was never able to reach a person to get the address and it was not visible on the house. Also due to city address improvements, some addresses on units may be updated by the time of PI. Therefore, we allowed the PI interviewer to update the address, if they were sure they were at the right unit but some part of the address needed to be updated. This allowed us to better match the people and refer to the right address during the interview. Stateside, 1,653 cases (0.92 percent) had their address updated and 412 cases (5.15 percent) had their addresses updated in Puerto Rico.

In Puerto Rico, the main change was usually just the correction of unit number, followed by updating just the house number, but there were almost all combinations of updating any of the address parts. Overall, the house number, unit designator, and physical description were changed about 25 percent (102-107 cases) of the time that they updated the address.

For Stateside, the main updates were correcting just the house number followed by updating the street name only. Again, almost all combinations of updating the address parts occurred. Most interesting is that the house number was updated 48.15 percent (796 cases) of the time that they updated an address. This is concerning since this is not

[22] Two CRs were immediately submitted and approved. On the RI processing side, TMO added a procedure before loading the RI database to check for these characters in all open text fields and replaced them with the English character if possible or else deleted any letters with these characters applied. In Post Processing, we did a very similar process except the letters with characters were mapped to their corresponding English version of the letter to be replaced.

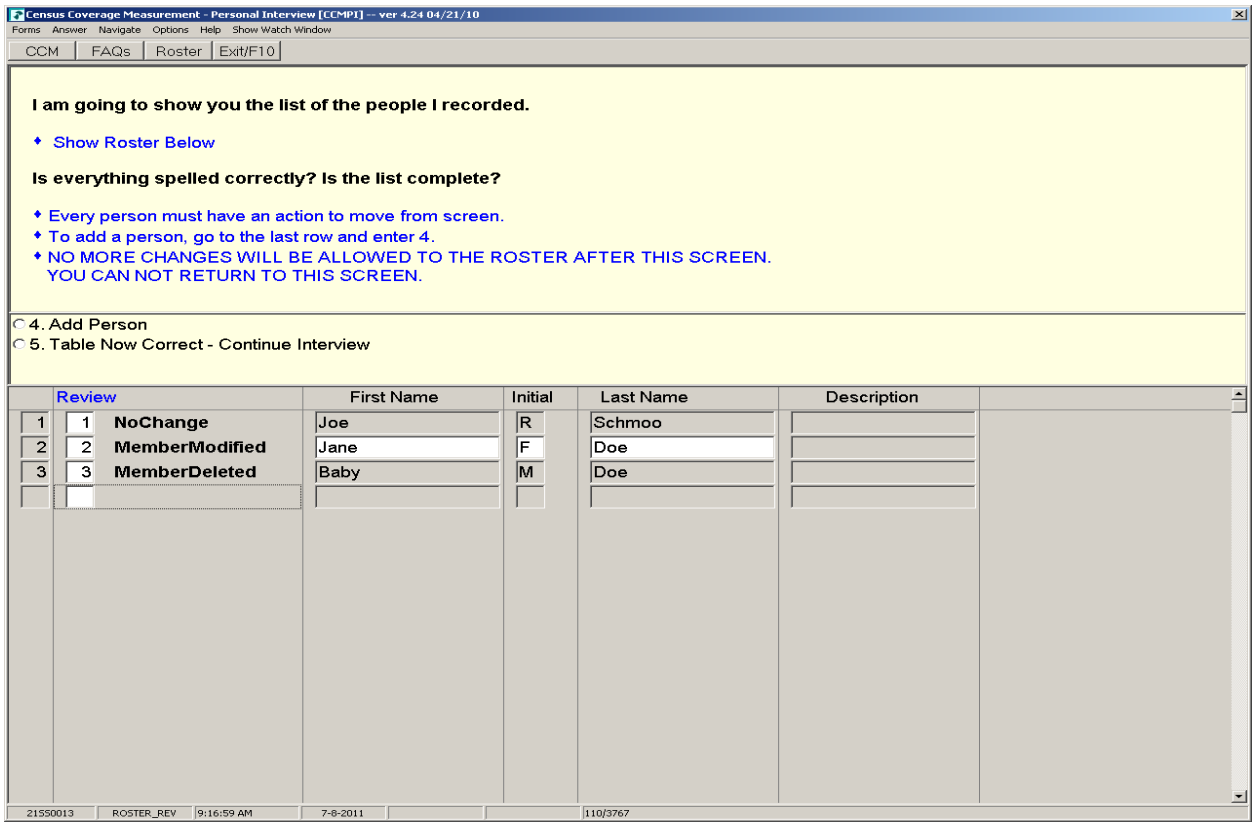
something that should change often and there is no way without an observation in the field to make sure they were truly at the sample address and that those were good corrections. While addresses were hardly updated, we need to revisit this issue in the future to maybe limit the type of address corrections allowed or set up some sort of verification procedure with a supervisor.

5.17.2 Review Screens – Person and Address

As a way to make the interview flow as smoothly as possible, there were two places in the instrument that stop the interview to review what was collected. This occurred at the end of any roster collection and also to review alternate addresses before asking about timing of when the people rostered lived at each address. It turned out that the person review screen became even more important than planned. In order for the instrument to work properly, it was needed to put a locking point in the software at the review screen where the interviewer could not go back and change the roster once passing that screen.

In the review screen for people, the interviewer had five choices to clean up the roster. They could confirm the people were listed correctly, correct the spelling of a name, delete someone already listed, or add a person. The last choice was to confirm that the entire roster was now correct. Figure 7 shows an example of the review screen for people with the first three actions entered and the last two actions as options.

Figure 7: Person Interview Instrument Person Review Screen



Overall, only 4.88 percent (20,689) of people rostered had their name edited in the review screen. Only 0.07 percent (314) of people were added and 0.07 percent (302) of people were deleted at the time of review. This screen seemed to work well with few complaints about the screen, but we did get complaints of wanting to go back and add or change the roster after getting past this screen. Therefore, we need to look at what are the triggers later in the instrument that makes a respondent think of someone after leaving the roster area when we spend so much time on the roster at the beginning. We plan to use paradata for later analysis on this issue.

Reviewing the alternate address questions connected to a person actually requires two screens. The first screen is a verify screen that lists all people who have no alternate address. On this screen, the interviewer can only confirm these people's only address is the sample address or add an address if necessary. The other screen is a review screen for people who have more than one address. The screen is person-based and lists all addresses collected so far for that person. (See Attachment A for exact wording.) Originally, we intended for this review screen to allow interviewers to only confirm or add an address for each person, but interviewers requested the ability to also delete an address. A total of 78,234 people went through this review screen. For 1.49 percent (1,167) of those people, respondents needed to add an address and 1.07 percent (838) deleted an address. An interesting finding was that 15.27 percent (128 cases) of the time, they deleted an address and then added a new address back either in the Census Day or Interview Day sections immediately following the review screens. It is possible this may have been the way the interviewers worked around the need to edit an address without going back in the instrument to where they originally collected the address.

5.17.3 Move Question in Review Section

As described in Section 5.12, we collected move dates to determine where people were on Census Day. While we had the two main move sections at the beginning of the interview, we also had a move section after all the alternate addresses were collected. If a respondent reported an alternate address in one of the general sections (such as Job or Second Home), we checked if they moved from this address or went back and forth. If they indicated that they moved, then we collected a date and treated this person now as an Inmover; a person who had moved **into** the sample address. There was a general assumption that people who indicated a move at this point in the interview must be Inmovers since they were listed as people who currently live or stay at the sample address.

While reviewing people who indicated a move in this section during clerical matching, there were some people that seemed to have contradictory information. By reviewing notes and answers to other questions, we found that some of the people who indicated a move in this section were actually not inmovers, but outmovers. This has two effects. First, it shows that some outmovers were listed on the main roster, which we did not expect, and also that the move question may not be indicating the right type of move. Through review of the responses to the "live here now" question, we estimate that around eight percent of the people who told us they moved in this section were really moving out of the sample address and not moving in. It is most likely that these situations were caught in matching and corrected. Overall, 4.88 percent (225) of people who indicated a

move got marked as an outmover as their final residence code. If the move question is used again in the future, it should include a screen/question to check if the person moved into or out of the sample address.

If we look at all people listed on the main roster (including probes), only 0.31 percent (1,247) of the people listed there are outmovers. Therefore, it is a limited issue. In addition, 0.48 percent (77) of people listed as outmovers ended up being coded as inmovers. Further research is needed to see why people would list outmovers on the main roster but it is not a large problem. We essentially need to confirm, rather than make any assumptions on questions related to people moving in or out of the sample address.

5.17.4 Probes for Tenuously Attached People

As discussed in Section 5.9, some of the people listed in the interview are added through probes. These probes were not well liked by either the respondents or the interviewers. They were considered repetitive and tedious and often interviewers and respondents will shorten or cut off the questions. We still believe that we need to do some probing to get the respondent to list people they may not consider part of the household, but we need to further review how we probe in the future to maybe a level that is more accepted by the interviewers and respondents. To better understand this issue, Table 46 shows some of the key characteristics for people from each of the different probes.

Table 46: Review of Person Interview Person Characteristics by Probe Where Person was Included†

Row	Characteristics of People	All People		Probe 1: Stays often		Probe 2: No other place to live		Probe 3: Missing children		Probe 4: Missing relatives		Roster Review		All Probes	
		Count	Percent/ Age Range	Count	Percent/ Age Range	Count	Percent/ Age Range	Count	Percent/ Age Range	Count	Percent/ Age Range	Count	Percent /Age Range	Count	Percent /Age Range
1	People that have no alternate address	343,652	81.20	545	6.89	325	30.72	490	48.51	182	37.22	189	68.23	1,731	16.11
2	People that are Non-White	143,451	33.89	3,075	38.88	483	45.65	546	54.06	277	56.65	134	48.38	4,515	42.03
3	People that are Hispanic	81,162	19.82	1,600	20.23	226	21.36	265	26.24	134	27.40	72	25.99	2,297	21.38
4	Peak Age Range for People	N/A	None	2,430	15-24	264	20-29	744	0-9	102	20-29*	73	0-9	2880	20-29
5	People counted at Sample Address at end of PI	334,159	78.95	0	0.00	315	29.77	426	42.18	162	33.13	174	62.82	1,077	10.03
6	People counted at Sample Address at end of CCM Person Matching	361,489	85.41	2,487	31.45	394	37.24	480	47.52	197	40.29	186	67.14	3,744	34.85
	Total People	423,242	100.00	7,909	1.87	1,058	0.25	1,010	0.24	489	0.12	277	0.07	10,743	2.54

Source: Person Interview Final Version Output and Results of CCM Final Person Matching

* Probe 4's peak age range contains only 20.86 percent of the people while all other peak age ranges contain at least 25 percent of the people.

† People are not unique to a row and people can be counted in more than one row. Hence, columns are not cumulative.

Row 1 shows the percent of people collected from the various probes that did not provide an alternate address. Some of the people from the probes may not have been listed in the main roster due to not understanding who should be listed. However, people added at some of the probes may also result from the people being tenuously attached to the household and likely having another address they stayed at. Therefore, we would expect the percent reported in Row 1 for people having no alternate address to be lower than for overall people, since most tenuously attached people should have another address. Specifically, the percent for Probe 1 is much lower than for the other probes since it specifically refers to people who have someplace else to live or stay. Probes 3 and 4 and the Roster Review seem to be still high percentages even though they are lower than the overall percent for all people reported. Some more research into people listed in the probes that have no other addresses should be done to see why they were not listed in the main roster.

Row 2 is the percent of people added by the probes that are non-white and Row 3 is the percent of people who are Hispanic. Typically, minorities have a more complex household structure. (Martin 2007 and Schwede et. al. 2005) We would expect that people added by the probes would be added during these probes given these complex household structures. Therefore, it makes sense the people added by the probes have a higher percentage of minorities than when all people are considered.

Row 4 looks at the peak age range for each probe. One of the populations typically missed within the Census is young males. (Martin 2007) We would expect that each of the probes would collect this age range (except for Probe 3 that was purposely targeting babies and children). In this case, the probes did more often collect the people we were targeting. Interestingly, when looking at all the children rostered (people age 17 and below), only 0.86 percent (874 people) were added through Probe 3. This may be a sign that underreporting of babies may not have been as big an issue for PI as it has been historically for the Census. In general, Rows 1 through 4 show the probes are collecting the type of people we are most concerned about.

The overall goal of the probes is to find people who should be counted at the sample address but at the same time it is acceptable to add people that would not be counted at the address in the end. As long as some of the people are counted at the sample address then the probe has some value. Row 5 shows how the person was coded at the end of PI post-processing. There was not one person added from Probe 1 that was listed automatically as being counted at the sample address. This is concerning because either we are listing the wrong types of people or the respondents are for some reason not capable of answering the necessary questions clearly. Row 6 is the final residence status the people added by the probes received in CCM after all CCM PFU and matching operations. We can see that the person matching and followup activities were able to increase the percent of people added by the probes counted at the sample address. In particular, these operations confirmed residence status at the sample address for 31.45 percent of the people from Probe 1 that the PI post processing could not confirm. As one

can see, 34.85 percent of all people added by the probe should have been counted at the sample address.

5.17.5 Whole Household Outmover Status and Roster Collection

In 2010, PI not only collected information on individual outmovers but also attempted to collect information for when an entire household had moved out of the Sample Address since Census Day. Since the entire household switched, the respondent would always be acting as a proxy since even if we were talking to the current occupant they did not live with that household on Census Day. In this scenario, the interview was set up to collect this information on the outmover household after collecting the current housing unit status and people (i.e., after the main interview). This section did not need to be completed for a case to be considered a sufficient partial. If the current resident did not know anything about the WHO, then the interviewer did not need to find another respondent and was instructed to end the interview.

An interviewer enters the WHO section in the following scenarios:

- The Sample Address is currently not a housing unit.
- The Sample Address is currently vacant.
- All people rostered at the time of interview have moved in after Census Day.

The WHO section first determines the status of the sample address around Census Day. If it was occupied, then the move date, roster, and the address the people moved to is collected if the respondent knows it. After those fields are collected, a copy of the main interview is conducted for the WHO using almost identical questions as we did for the Sample Address Roster.

The WHO section can more than double the length of the interview and it is usually difficult to find a knowledgeable respondent. A knowledgeable respondent knows the names, demographic characteristics and move dates for the outmovers. This section also needs its own set of procedures, development, and testing. We need to review the costs and benefits of including the WHO section in future applications to make sure we are using resources efficiently.

A total of 34,861 interviews went into the WHO section, 18.67 percent of all cases. Of those, 9,949 (28.52 percent) were due to a whole house of inmovers. In general, respondents could supply the housing unit status of the unit on Census Day first. The missing data rate on status was only 12.25 percent (4,270 cases). Most of the time the interview went to the WHO section (67.13 percent (23,403 cases)), the sample address was vacant or not a housing unit on Census Day. For 20.62 percent (7,188) of the units, the respondent indicated that it was occupied on Census Day; these are the cases that are true WHO cases. The next question is a screener question to see if the respondent knew the names and ages of the people who lived there. Only 45.44 percent (3,266) of respondents felt they could provide this information; so we collected information on WHO households in only 1.75 percent of all cases.

A total of 6,745 WHO people were rostered. The smallest roster was one person and the maximum roster was 14 with almost half (47.06 percent) of the rosters being just one person. This may be the respondent telling us about only one of the people in the household that they may have come in contact with or see mail for.

Table 47 shows some of the data and the final residence status code of the people collected in the WHO section.

Table 47: Person Interview Characteristics of Whole Household of Outmovers People†

		All People		Whole Household of Outmovers People	
		Count	Percent	Count	Percent
1	Percent of People With Full Name	413,455	97.69	5,371	79.63
2	Percent of People with Age Reported	394,832	93.29	4,512	66.89
3	Percent of People Knew if Move was Before or After April1	50,905/ 51,839*	98.20	6,615	98.07
4	Percent of Addresses collected that are complete	30,753/ 60,590*	50.76	863/4,54 1*	19.00
5	Percent of houses that are unrelated people	60127/ 158,024*	38.05	1,950/ 3,373*	57.81
6	Percent counted at Sample Address at end of CCM	361,489	85.41	4,445	65.90
7	Percent Unresolved at the end of CCM	10,785	2.55	1,944	28.82
	Total People	423,242	100.00	6,745	1.59
Source: Person Interview Final Version Output and Results of Final Person Matching					

† People are not unique and people can be counted in more than one row. Hence, columns are not cumulative.

*In count, if the denominator for the percent is not the entire analysis universe then the denominator is listed in addition to the count.

Rows 1 and 3 show that respondents who said they knew the WHO people were able to give us names and move dates most of the time, but were not able to give us age as consistently (Row 2). They struggled with giving us a complete address (Row 4).

All WHO people were sent to clerical review in the CCM Person Matching Operation. Of those, 41.85 percent (2,823 people) went out to person followup and 20.13 percent (1,358 people) more may have needed to go to followup to be resolved but could not

because the names were not complete. In the end, 28.82 percent (1,944) of the WHO people listed were unresolved at the completion of all CCM operations and we only got enough data to assign a residence status code to 4,801 WHO people (71.19 percent).

5.17.6 Contacting the Respondent – Front Section

The first section in the instrument (referred to as the "Front Section") is responsible for assisting the interviewer in confirming they are at the correct address and finding the best person to be the respondent. There are 26 different questions following many paths to try to cover all possible situations the interviewer may encounter at the door. (See Attachment A for the simplified path and wording of a standard interview). Through debriefings, observations, and data review, it appears the front worked most of the time, but many of the data issues we encountered during review, such as miscodes of outcome code and attempt type, came from the front. It appears that the front section may have been too scripted and forced the interviewer to force a "fluid" conversation into a set of overly restrictive answers. The following are a few examples:

- One of the problems created a loop that the instrument could not handle and closed the case without it ever being worked. We needed to check for these types of cases and resend them out daily.
- The main issue most interviewers reported was that they needed more guidance and help on getting a person to do the interview for soft refusals[23]. They felt they could not engage the respondent enough to convert them. While we did some training on this, more training would help, but one of the issues could be that, because the front was so scripted, the interviewers were more focused on entering the correct answers in the proper screen than engaging with the respondent.
- When reviewing the trace files (keystroke output from the BLAISE instrument), we found the interviewer often went back and forth within the front section at least once, if not more.

Alternatively, a simplified front section focusing on reminding/checking that the interviewer follows the rules to obtain a knowledgeable respondent may be better. Extensive testing would need to be done to be sure this process works. We recommend reviewing the opening screens of the demographic current surveys to see what approach they take.

5.18 What impact did the Nonsampling Error Reduction Initiatives have on the Person Interview operation?

In September 2009, the Census Bureau announced an initiative to reduce nonsampling error in the CCM program. In order to make the change cost neutral, CCM managers reduced the sample size for operations *after* the CCM Independent Listing, which was almost complete

[23] A soft refusal is considered someone who is unwilling to do the interview but did not directly refuse. They may say they are reluctant respondents that usually have concerns about doing the interview or about the government. We call these people soft refusals because often with a little discussion and explanation of the survey they are willing to do the interview.

nationwide, and diverted the resulting funds towards approaches to reduce the nonsampling error. We list below the initiatives that were implemented for PI and the results:

- **Higher fieldwork reinterview rates** – We increased the RI random sampling rate from 8.3 percent to 12.5 percent. This helped ensure that we had a higher quality level allowing us a better chance of catching falsified PI cases. There is no way to measure if the higher sampling rate led to higher quality. However, even with the higher RI rate, we encountered an issue as a result of reducing the PI sample size, which still lowered the overall RI workload. Most RI interviewers complained about not receiving enough work to keep them busy.

- **Adding training modules to interviewer training** – We added several modules to interviewer training, including localized training scenarios, training on situations due to current economic conditions (temporary movers, etc.), and training on probing for alternate addresses. Observers and interviewers both commented that they did not have as good a grasp on content of the interview as they did on the automation when they started working. They felt that the hands-on scenarios were very valuable in training and would have liked more training on handling refusals and noninterviews. There were no comments on any of the areas where we added more training. We feel this means they did not have any issues with those. The localized training could have been more focused to smaller areas within the region instead of for the whole RCC, such as covering how to handle seasonal vacants for training sessions containing interviewers that would be working in beachfront communities.

- **Smaller employee-to-supervisor ratios** – The intent was to have fewer lower-level field staff assigned to each supervisor and ensure a greater control over the quality of the field work by allowing more monitoring of work at each level. The regions did confirm that for PI this was very valuable. While the CLs did not have to handle paper tracking, they did have to help with automation issues and with more complicated procedural issues of interviewing (help with contact, how to explain the interview, etc.). They also worked as interviewers where they were invaluable in handling the hard to contact cases. FOSs actually ended up working more hours than expected, but mainly handling payroll. Assuming payroll is automated in the future; some RCCs suggested it might be more appropriate in the future to have one or two more regional technicians assigned and not have FOSs.

- **Extending person reinterview for one week** – Due to scheduling constraints, the original plan was to stop sampling for RI one week before production finished, to allow time to complete the reinterview cases. With a smaller workload, we were able to reduce the duration of some of the following activities thus allowing the week extension for RI. In the last week, 385 more RI cases were selected to be reviewed. Most (135) were selected because of Hard Failing an interviewer. If this extra week had not been added, these cases would not have been reworked.

- **Completing an Extra Observation on all interviewers –** Common practice is for a CL to observe each interviewer or reinterviewer within a week of them starting work to make

sure they are following procedures and to help them with any areas where they have questions. After the initial observations, it is uncommon for interviewers to ever be observed again. To make sure procedures are followed all the way through the survey, an extra observation was added later (3 to 4 weeks after interviewers started work) in the survey period. This appears to have been a possible hindrance to the survey. It became very cumbersome to make sure everyone had the extra observation and took up much of CLs and FOSs time, as well as office time in tracking the progress. It was also an issue that some people no longer had a workload high enough to meet the requirements of completing an observation, since at this point they were just trying to convert the hard to reach respondents. A total of 4,412 extra observations were done or attempted to be done. Of those only 10 interviewers/reinterviewers were marked unsatisfactory. We do not think it adversely affected quality, but we do not believe data quality was improved either when considering the time, cost, and effort needed to accomplish the task. See the CCM Person Interview Quality Profile Early Closeout Draft for more information on interview observations.

Overall, the nonsampling error reduction initiatives are very difficult to analyze for specific quality results, but the common understanding of most stakeholders is that they helped ensure quality and we did not see any negative impacts from the changes. Some of the issues with staff complaints about not having sufficient work at times are likely a result of the short timing between the transition of starting staffing and training for a much higher workload and the actual decreased workload. Assuming these procedures would be built into the original design in the future, staffing assumptions would be based on these procedures and lack of sufficient workload should not be an issue.

6. Related Evaluations, Experiments, and/or Assessments

The following Assessments/Evaluations are being conducted about 2010 PI or they use data collected during 2010 PI.

- PI Behavior Coding Assessment
- PI Debriefing Experiment
- Comparative Ethnographic Studies of Enumeration Methods and Coverage in Race/Ethnic Groups
- 2010 Census Coverage Measurement Person Interview Quality Profile Early Closeout Draft

The following is the list of assessments on all other CCM operations:

- Assessment for the 2010 Census Coverage Measurement Sample Design
- Assessment for the 2010 Census Coverage Measurement Initial Housing Unit Independent Listing, Matching and Followup Operations
- Assessment for the 2010 Census Coverage Measurement Person Matching and Followup Operations

- Assessment for the 2010 Census Coverage Measurement Final Housing Unit Matching and Followup Operations

7. Lessons Learned, Conclusions, and Recommendations

7.1 Lessons Learned

There were many lessons learned throughout the operation. The PI team met and developed a set of lessons learned from the headquarters perspective (Linse, 2010). In addition, Field Headquarters staff created a summary of the CCM PI Debriefing of FOSs, CLs, interviewers and reinterviewers to create a summary of insights and suggestions from the field staff that can act as their lessons learned document. There was another document of a CCM Debriefing on PI operations conducted by Regional Managers in May 2011. This acts as lessons learned from within the Regions. The following are a few of the key lessons learned from these documents in no particular order:

- Many respondents prefer to finish an interview by phone after initial contact. Both the RCCs and the interviewers conducted phone interviews when the respondent called back responding from a Notice of Visit left at the address. This is what most respondents who contacted the RCC or interviewer requested.
- Training for automation use (i.e., laptop functions, how to open and run Case Management) can be centralized or done through the laptop (Computer Based Training). Most training on use of laptop, payroll, and standard functionality does not require that a trainer be in the room. These could be easily conducted through a training application on the laptop or through connecting with a trainer over the laptop. While it may not be practical for all training to be this way (some component of classroom training would still be necessary), a cost reduction may result by using this alternative when feasible.
- The main issue interviewers had was getting respondent cooperation. They felt they were not able to connect to reluctant respondents and the training did not prepare them on how to respond to the questioning/hostile respondents they encountered during the operation to know what to say to get cooperation. We need to add/change training on getting cooperation and make "real practice" scenarios so interviewers are more comfortable and are able to handle these types of situations from the beginning.
- Proxy Guidelines were difficult to follow not only for the interviewer but also for some of the regional staff. The guidelines need to be more consistent and have fewer exceptions, or be completely automated into the instrument.
- Refresher Training was one of the most useful parts of overall training.
- Contacting and getting a respondent to cooperate is a more fluid concept and the automated instrument may have been too rigid in collecting all the contact data at the beginning, limiting the interaction with the respondent.
- Interviewers and Respondents will change the answers and back up in the instrument. More testing and functionality is needed to keep a high level of data quality when this happens.
- In CMOCS (or whatever tracking system is used in future), certain PI data need to be always retained such as notes, count attempts, and case history, even when RCC staff wants an interviewer to start an interview over as new.
- Tracking of workload and progress varies region to region. Allow a more interactive system to develop reports within CMOCS to limit the need to create unique reports separately in a different format.
- All systems should be automated to remove any paper tracking.

- Serious respondent fatigue was occurring by the time PI took place. We need to review to see if there are ways to limit the contacting of people between all of the CCM operations (plus the possible Census contacts).

7.2 Conclusions

Overall, the CCM PI and RI operations were a great success. PI had an overall response rate of 98.4 percent. The missing data rate was under 3.0 percent for all of the residence questions. For all the data we collected, we knew if the person should be counted at the Census Day residence for over 98 percent of them. We collected 60,950 alternate addresses in PI, and only 7.74 percent of the addresses did not provide any data to establish a location. After all work was completed, the 2010 CCM PI operation was under budget by $8,333,640 (25.98 percent). We feel automating PI was a big success and the benefits far outweighing any issues.

We found the following sections of the instrument worked particularly well:

- The probe sections to find more people that should have been counted at the sample address picked up people that ended up being counted at the address.
- The whole household outmover section did very well at collecting the housing unit status and picked up some people who had moved out.
- Every section collected some addresses and no complaints were reported by interviewers about the number of times we ask about various addresses.

There are several small areas where the wording of the questions in the instrument could possibly be fine-tuned in the future:

- The front section was too structured for the times when finding a respondent was difficult or unusual.
- The probes for more people worked in getting more people but were seen as a nuisance by interviewers and seemed repetitive to respondents. This did not give a good impression early in the interview.
- The collection of move dates along with the cycle questions to people with various addresses did not always seem clear and straight forward to respondents.
- The interviewers struggled with the concept of when they could use a proxy and what constituted a knowledgeable proxy in the different situations. In the end, proxies could not provide much data and they had higher missing data rates on all fields. We need to review the benefit of proxies with newer estimation procedures due to the cost and complexity of using them. If we continue to use them, we need to find a better way to guide the interviewer in the proxy interviewing process.

Overall, only 18.43 percent (78,023) of people needed to be clerically coded to a final residence status at the end of PI, which was manageable.

7.3 Recommendations

The major recommendations through either lessons learned or through the results presented in this assessment are:

1. Add a telephone phase or self-response phase to PI to get earlier responses from cooperative respondents.
2. Revisit the timing for PI, i.e., move it as close to Census Day as possible. This should help reduce respondent reluctance to participate (they will not think Census is over) and should lead to higher data quality by reducing the number of people who will have a change in their living situation since Census Day.
3. Training not only needs hands-on practice interviews for learning the instrument and automation, but needs realistic difficult interaction scenarios to better prepare our interviewers for gaining cooperation at the door. If the sample continues to be small, consider using demographic or economic interviewers from other Census Bureau operations who already have this ability and will not need as much training.
4. Testing of any instrument or system needs to be done in a realistic setting as well as the usual structured one. This allows not only planned scenarios to be tested, but also for unexpected situations to come to life. This could be done through small "live" field tests or by bringing to the testing sessions experienced field staff that understand and can share what field situations arise and can replicate them on the laptop. Timing of testing needs to allow for change to procedures and training before production and not just with the start of production.
5. Research how to implement the necessary roster probes while limiting their repetitiveness.
6. Review and modify the section that collects the times of staying at various individual addresses, and ensure confirmation if the person is moving in or out of the Sample Address.
7. Review keystroke tracking for any indication of where the interviewer/respondent paused, where an interviewer started to go back, and where answers were most often changed to determine which questions are causing confusion or acting as memory triggers to change other answers.
8. Review the process for creation of a Spanish translation. We all agree one is needed for use in the instrument, but there should be a way to smooth the development and make more universal translations that are easier to use in a conversational setting. All translation should at least be done by one entity and not various sources.
9. Review all proxy results and costs to determine if proxies should be allowed in the future or if process of using proxies should be modified.
10. Allow time in the schedule for detailed data review before beginning of post processing. Even with the best testing, it is likely we will get some data that are unexpected. It is best to allow time for those situations to be reviewed and corrected. Also have a contingency plan to remove cases/data from production processing to limit the amount of cases forwarded through production after a later version has been received.
11. With the changes to CCM through automation of other operations and possible changing to multi-mode for CCM PI, employee ratios and QC rates would have to be completely reviewed and tested.

8. Acknowledgements

I would like to acknowledge all the hard work that the entire CCM PI and PI QC Team put into not only this assessment, but also into making the entire CCM PI a success. I would like to particularly thank those that reviewed this assessment and added their own insights.

Special thanks to Elizabeth Nichols and Jennifer Hunter Childs for not only being leading forces in the development of the PI questionnaire, but for taking the time to give me all the background in PI instrument development since 2000. Your insights into this review were truly valuable.

Thank you to John Jones, Jennifer Weitzel, and Kopen Henderson for their assistance in checking this assessment for accuracy.

9. References

Adams and Nichols (2007). "Census Day Residence Evaluation," 2006 Census Test Evaluation 5, April 27, 2007.

Hartman, Patricia (2012). "2010 Census Coverage Measurement Person Interview Quality Profile Early Closeout Draft," K:\2010 CPEX Canceled Studies\Quality Profiles\Person Interview\Person Interview QP_early closeout-draft after ESC Briefing - 1-27-2012.pdf, January, 26, 2012.

Hunter and Nichols(2004). "Final Report from Westat: Cognitive Testing of the Proposed 2006 Census Coverage Measurement Person Interview," DSSD Census Coverage Measurement Memorandum Series #2006-D7-02, November 10, 2004.

Linse and Stone (2010). "Comparisons and Explanations of Differences in Person Interview Noninterview Rates from the 2000 Accuracy and Coverage Evaluation and the 2010 Census Coverage Measurement Program," DSSD 2010 Census Coverage Measurement Memorandum Series #2010-I-P-02, December 13, 2010.

Linse, Kyra (2011). "Lessons Learned for 2010 Person Interview Operation," DSSD 2010 Census Coverage Measurement Program Series #2010-D7-12, DRAFT.

Martin, Elizabeth(2007). "Strength of Attachment: Survey Coverage of People with Tenuous Ties to Residences," Demography: Vol. 44. No2, May, 2007.

Schwede, Blumberg, and Chan (2005). "Complex Ethnic Households in America," Oxford, UK.

"Standard Definition: Final Dispositions of Case Codes and Outcome Rates for Surveys," The American Association for Public Opinion Research Publication, Revised 2011 Edition.

Stuart, Leslie, and Brady (2003). "Assessment Report for A.C.E. Field Operations," Census 2000 Informational Memorandum #132, February 5, 2003.

Ramos, Magdalena (2010). "Summary of the 2010 Census Coverage Measurement Person Interview Instrument," DSSD 2010 Census Coverage Measurement Memorandum Series #2010-D8-04.

Whitford, David (2001). "Accuracy and Coverage Evaluation Survey: Person Interviewing Results," DSSD Census 2000 Procedures and Operations Memorandum Series B-5*, February 28, 2001.

Whitford, David (2007). "Final Report for 2006 Census Test Evaluation #5: Census Day Residence," DSSD 2010 Census Coverage Measurement Memorandum Series #2006-B-05, September 24, 2007.

Whitford, David (2009). "Recommendation to Reduce Nonsampling Error in the 2010 Census Coverage Measurement Program," DSSD 2010 Census Coverage Measurement Memorandum Series A-43, December 7, 2009.

Attachment A: Simplified Review of Instrument Wording

(RAMOS, 2010)

Introduction

The purpose of the PI interview is to obtain information about the current residents of the sample-housing unit. This includes those who may have moved into the selected housing unit since Census Day (April 1, 2010). The instrument also collects data for those who moved out of the sample housing unit between Census Day and the time of the CCM interview (outmovers). For each person, the CCM PI collects the same demographic information as collected during the Census: including name, relationship, sex, age, date of birth, Hispanic origin, and race. The PI collects information to determine where each current resident was living on Census Day and where each outmover currently lives. The CCM PI also collects information to determine if there are any other addresses where the person may have been counted in the census and information necessary to geocode the alternate addresses. Other features of the instrument include:

- Separate scripts for nonproxy and proxy interviews.
- Separate scripts for in-person and telephone interviews.
- A separate full Spanish language translation. The instrument also records whether the interviewer used the Spanish or the English version.
- Special formatting for the collection of Puerto Rico addresses.
- Item-specific help and frequently asked questions.
- Automated fills within the questions and response categories (including name fills, address fills, and date fills).
- References to a visual aid to assist respondents in completing the interview.

There are 15 separate modules corresponding to separate portions of the interview in the CCM PI. Although these modules are labeled A-M, W, and N within the CCM PI specification, the labels are unknown to the respondent. Depending on the data collected, only the appropriate questions in each module are displayed for the interviewer to read to the respondent.

One of the features of the automated instrument is to present each question so that it can be read to the respondent verbatim, whether that is in-person face-to-face interviews, or on the telephone, and whether the interviewer is talking to a current household occupant, or a proxy respondent. The question text also changes depending upon whether the question pertains to the respondent (where the question text uses "you") or another occupant of the housing unit (where the question text will use that person's name). For ease of reading within this document, we present only the text assuming that the interviewer is talking to a current household resident during an in-person interview, and the questions refers to the respondent. Bolded text within the document signifies what is read to the respondent. Text in all capitals refers to text that is filled with a name, address or date. Included next is a summary of each module's goal and an overview of the main questions used within each module.

MODULE A: INTERVIEW ATTEMPT

The goal of this module is to collect information used for the subsequent question wording fills (e.g., proxy or nonproxy respondent; personal visit or telephone interview) and to also collect information on attempts at an interview. The screens in this module include a set of questions and instructions designed for response by the interviewer and a set of questions designed for response by the household (or proxy) respondent.

Answered by interviewer:
There are screens requiring the interviewer to enter the following information:

1. Type of contact: Via Personal Interview or Telephone, and if at sample address or at a proxy.

2. Whether or not the interviewer located the unit.

3. Result of contact attempt. If a personal visit, someone answers door, no one answers the door, it is a vacant unit, the unit is not a housing unit, or the unit is inaccessible. If a phone interview, someone answers or they do not (rings but no answer, busy signal, answering machine, out of service number, number change, or pager/FAX).

4. If a proxy interview is being attempted, the interviewer must record the reason for a proxy interview and whether or not a knowledgeable proxy is available.

5. If trying to confirm if vacant or not a housing unit, the interviewer must record if they found a proxy respondent who can 1) confirm the current status of the unit, and 2) the status of the unit on Census Day.

6. After six attempts, indicate if they want to confirm a unit as vacant or not a housing unit by observation.

Answered by respondent:

The objectives of these questions are to find a knowledgeable respondent and to verify the address location.

The introduction read to the respondent is, "**Hello, I'm (YOUR NAME) from the U.S. Census Bureau. (Here is my identification.**)"

The following questions are asked after the interviewer has shown identification and introduced themselves to the respondent:

(For personal visit) **Is this SAMPLE ADDRESS?** (For telephone interview) **Have I reached HH PHONE NUMBER/PRX PHONE NUMBER? ;** where HH means Household and PRX means proxy. "SAMPLE ADDRESS" is used in this document in place of the actual sample address that would be displayed in the instrument for the particular case.

(If Interviewer is trying to find out if the unit is vacant or occupied) **I have some questions about SAMPLE ADDRESS. Does someone live/ does anyone currently live at SAMPLE ADDRESS?**

The survey introduction is for an in-person interview**: "As part of the census, we are contacting households to make sure we counted everyone correctly. I have some questions about (this address/SAMPLE ADDRESS.)"**
The following questions are asked after the interviewer has introduced the survey to the respondent and determined if the respondent is knowledgeable.
(If Interviewer is speaking with someone at the sample address) **Do you currently live or stay here?** (And, if necessary) **May I speak with someone who currently lives or stays here?**

(When speaking with a proxy, the PI poses a question to determine whether the proxy is a knowledgeable respondent for an occupied unit and the interview can continue.**) I am going to be asking questions about the people who live at SAMPLE ADDRESS – things like names, ages, and where they lived on April 1st. Do you know that information?**

(When determining whether an interview should be conducted at a seasonal address**) Is there anyone staying here now whose stay will last for more than two months? Does everyone staying here have another place where they usually live?**

(Determining how to refer to the unit within the questionnaire) **Is SAMPLE ADDRESS a house, townhouse, condominium, apartment, mobile home, or another type of place?** (If the address from our records was only a physical description**) I need to refer to SAMPLE ADDRESS during the interview. How would you like me to refer to it?**

MODULE B: ROSTERS

The goal of the module is to collect names of all current residents, which includes any people who stay overnight at the unit more often than they stay anywhere else, and any other people who might have been missed in the census. This latter group includes people between residences, highly mobile people, nonrelatives, extended relatives, and children. There are two different paths for creating a list of people living at the unit. The first path is for people who are staying at a seasonal unit or in transitory living quarters who have no other place they usually live. We list that first here because it is shorter. The second path is for the more typical permanent housing unit. Most CCM rosters will be created using the questions from the second path.

Path 1: This path is used when there are people at a seasonal unit or when respondent indicates some people are staying at the sample address temporarily but have no other place they usually live (in transitory living situation). We expect this path will not be used often.

What are the names of the people who are staying here and have no other place where they usually live?

First name? Middle initial? Last name?
Anyone else? (What is that person's first name? Middle initial? Last name?)

Path 2: This path is used when the sample address is not a seasonal home. This is the typical roster question path and most rosters will be created using these questions. There are five main questions and then a verification of the names collected, with the ability to add and delete names at that screen. As a reminder, the question text presented here assumes a non-proxy interview; there is alternate wording for proxy interviews. If the answer to any of the five main roster questions is yes, then the respondent is probed with, "What is that person's/childs's first name? Middle initial? Last name? Anyone else?"

1. We'll start by making a list of everyone who lives or stays here now. Let's start with you. What is your first name? Middle initial? Last name?
Anyone else? (What is that person's first name? Middle initial? Last name?)

2. Is there anyone who has another place to live but stays here often?

3. Is there anyone who is staying here until they find a place to live?

4. Are there any babies, foster children, or other children who stay here that you did not mention yet?

5. Have I missed any relatives or unrelated people who live or stay here?

After collecting the names, there is a verification of the roster. At this screen, names can be edited, added, or deleted.

I am going to show you the list of people I recorded. Is everything spelled correctly? Is the list complete?

If anyone is identified with the second roster probe which asked about people who have another place to live, the interviewer will collect the other place where that person stays using the standard set of address questions developed for the PI. The answers to these questions are used to geocode the address and find it in our records. The questions were developed based on what the estimation team and the Census Bureau analysts need in order to be confident that they are looking at the correct address within our records. This set of questions is used throughout the instrument.

Standard set of alternate address questions:

1. What is the address of the other place NAME stayed? (House Number, Street name, Unit Designation, City, State, Zip, and Country? collected in appropriate fields)

1a. (If the respondent says 'I don't know') **Do you know the city, state, or any other part of the address?**

1b. (If only city and state are given and they are the same as the sample address) **Is that place more than a mile away from SAMPLE ADDRESS?**

2. (If address is in the United States or Puerto Rico, and had at least a city name) **What are the cross streets closest to that place?**

3. (If address is in the United States or Puerto Rico, and had at least a city name) **Are there any landmarks nearby, such as schools or hospitals that would help someone find that address? If yes, please describe.**

4. (If address is in the United States or Puerto Rico, and had at least a city name) **What are the names of the neighbors who live near that place?**

MODULE C: IDENTIFYING INMOVERS

The goal of the module is to identify inmovers and to collect addresses where those people moved from. An inmover is someone who did not live at the sample address on Census Day, but does on interview day. The questions in the module are person-based. That means the entire set of questions is asked for each person before returning to the top of the module and repeating the questions for the next person on the roster. The following questions are asked about everyone on the roster. Answers to these questions determine whether each person was a Census Day inmover, and allows us to investigate whether a person was counted at both the inmover and the sample addresses, and where each person should have been counted on Census Day. These questions reference one of the visual aids – the calendar – that is presented on the Information Sheet (see attachment) .

Now using your calendar think back to where you were living on April 1. Were you living here on April 1 or somewhere else?

The following questions are asked of those determined to be inmovers (a "somewhere else" response to the previous question):

What was your address on April 1? Using standard set of address questions (see Standard set of alternate address questions in Module B)

Is the place you were staying on April 1 a house or apartment or another type of place shown on List A (on the information sheet I gave you)?

List A is a visual aid on the Information Sheet that displays a list of typical GQ such as dormitories, nursing homes, correctional facilities, etc. (Telephone interviews use an alternate wording that does not reference the visual aid.)

Is SAMPLE ADDRESS your only residence now, or do you still spend some time during the year at INMOVER ADDRESS?

(If the sample address is the only residence now) **What date did you move here?**

(If needed) **Did you move in before April 1, after April 1, or on April 1?**

(If the sample address is not the only residence now) **Please tell me how much time you spent at each address in the past year.**
Example of notes:

- Moved into this address in 2010; temporarily away during school year in Atlanta
- Only in Denver for a short visit during April; most of the time in Fayetteville
- M-F in L.A. during March & April for job; rest of time in Stockton

MODULE D: OUTMOVERS

The goal of this section is to identify any outmovers. Outmovers are people who lived at the sample address on Census Day, April 1st, 2010, and no longer live there. These are typically people who have moved out, but it also includes people who have passed away. There is only one roster probe used to identify outmovers. The CCM PI then collects the names of the outmovers, the date the person left the sample address, where the person moved to, and the type of place the person moved into. After these data have been collected, the list of outmover names is reviewed with the respondent. If the respondent indicates that he/she knows address information for them, all other modules collect information about outmovers as well as current occupants. The questions in this module are:

(Roster question used to collect outmovers): **Now let's talk about SAMPLE ADDRESS. Was there anyone else living or staying here during March or April who is no longer living here? What is that person's first name? Middle initial? Last name? Anyone else?**

(To determine if the person is an outmover): **What date did NAME leave SAMPLE ADDRESS to live somewhere else?**
(If needed) **Did NAME pass away/leave SAMPLE ADDRESS before April 1, after April 1, or on April 1?**

(To determine if the current respondent is a knowledgeable respondent for the outmover): **Do you know NAME well enough to answer questions about other places where they might have stayed during March or April?**
(If the respondent cannot provide this information, then they will not be asked any further questions about these people.)

The next question is the type of place question.
Is the place NAME was staying on April 1 a house or apartment or another type of place shown on List A (on the information sheet I gave you)?
What is the address of that place?
Using standard set of address questions (see Module B)

Review of outmover names: **I am going to show you the list of people who have (moved out/passed away/moved out or passed away).** Show list Names of outmovers
Have I spelled all names correctly?/Is that correct?

(This review functions exactly like the roster review located in Module B with edit, add and delete capabilities.)

MODULE E: DEMOGRAPHICS

In this module, the interviewer collects demographic characteristics about all listed people. These data facilitate matching the people collected in the CCM interview to people listed in the census for the sample housing unit. These questions are administered in a topic-based format. That means the same question (or topic) is asked for each person on the roster before going onto the next question (or topic). All of the questions, except for the nickname question (which is not a question in the census), were identical to the demographic questions used in the specification for the planned automated Census Nonresponse Followup. The demographics questions used in the interview are:

Nickname or other name: **Do you ever go by a nickname or a middle (or maiden) name?**
(If yes to nickname question) **What is the other name you go by?**

Tenure: **Do you or does someone in this household own this house with a mortgage or loan (including home equity loans), own it free and clear, rent it, or occupy it without having to pay rent?**

Reference person: **Of the people who live here, who owns/rents this house?**

Relationship: **Please look at List B on the handout I gave you at the beginning of the interview. How are you related to REFERENCE PERSON NAME?([24])**

1-Husband or wife
2-Biological son or daughter
3-Adopted son or daughter
4-Stepson or stepdaughter
5-Brother or sister
6-Father or mother
7-Grandchild
8-Parent-in-law
9-Son-in-law or daughter-in-law
10-Other relative
11-Roomer or boarder
12-Housemate or roommate
13-Unmarried partner
14-Other nonrelative

Sex: **Are you male or female?**

[24] The reference person is the person who owns or rents the housing unit as identified in the preceding question.

Date of Birth and Age: **What is your date of birth?**

(If date of birth is provided the instrument calculates age as of Census Day and verifies it with the respondent.) **For the Census, we need to record age as of April 1, 2010. So, just to confirm – you/NAME were/was CALCULATED AGE/less than one year old/not yet born on April 1, 2010?**

(If age was incorrect, allow respondent to correct date of birth.) **Since your age as of April 1, 2010 was AGE, can you help me correct your date of birth? I have <Date of birth>. What should it be?**

(If the date of birth is unknown) **What was your age on April 1, 2010?**

Relationship verification (used in cases where the relationship seems to be inverted, i.e., a "parent" appears younger than the child): **I have recorded that NAME is your parent/parent-in-law. Is that correct?**

The following questions on Hispanic origin and Race allow for multiple responses:

Hispanic origin: **Please look at List C. Are you Hispanic, Latino, or Spanish origin?**

(If the answer is yes) **Are you Mexican, Mexican American, or Chicano; Puerto Rican; Cuban; or of another Hispanic, Latino, or Spanish origin; for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on?**

Race: **Please look at List D and choose one or more races. For this census, Hispanic origins are not races. Are you White; Black, African-American or Negro; American Indian or Alaska Native; Asian; Native Hawaiian or Other Pacific Islander; or some other race?**

(If NAME is American Indian or Alaska Native) **You may list one or more tribes. What is your enrolled or principal tribe?**

(If NAME is Asian) **You may choose one or more Asian groups. Are you Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, or another Asian group, for example Hmong, Laotian, Thai, Pakistani, Cambodian, and so on?**

(If NAME is Native Hawaiian or Other Pacific Islander) **You may choose one or more Pacific Islander groups. Are you Native Hawaiian; Guamanian or Chamorro; Samoan; or another Pacific Islander group, for example Fijian, Tongan, and so on?**

(If NAME is some other race) **What is your other race group?**

MODULE F: ALTERNATE ADDRESS QUESTIONS

In this module, the interviewer collects any other addresses where the people on the roster should have been counted in the census, or might have been duplicated in the census. Internally, Census Bureau staff call these addresses, "alternate addresses." The types of alternate addresses for which probes are provided match the types of places which have historically caused census omissions and duplication, including college residences, second homes, work-related or military addresses and other relative's homes. Similar alternate address questions are also asked in the Coverage Followup (CFU) interview.

For each type of alternate address, the main question is asked for the entire household once. The ages of the people on the roster are used to populate the question text appropriately. For example, when we ask about any college residences, the question text only refers to those people of college age. Then, for anyone identified as having that type of address, person-based questions are asked to collect the alternate address. The PI question design does not assume that all of the people in a household with a college address would necessarily have the same college address.

The main alternate address questions follow. If an address is collected, a standard set of questions appropriate for each type of address is also asked. Many of those questions are identical or very similar to those asked in Module B. Additionally, for some types of alternate addresses an additional question was needed to determine whether the address was a housing unit according to census definitions.

Introduction to module: **Some people have more than one place to live or stay and could be counted in more than one place. The Census Bureau would like to make sure everyone you mentioned was only counted once.**

College alternate address:

During March or April, were you or was NAME attending college?

(If yes) **Who was attending college?**

What is the address where you were staying in March and April?

Other Relative address:

During March or April did you or NAME live or stay part of the time somewhere else with a parent, grandparent, son, daughter, or some other relative?

(If yes) **Who stayed somewhere else?**

What is the address of the other place you stayed?

Military address:

During March or April, were you or was NAME away because of military service?

(If yes) **Who was away because of military service?**

Were you gone for two weeks or less or for more than two weeks in March or April?

(If more than two weeks) **Were you staying in the US or outside the US?**

(If in US) **What is the address where you stayed?**

Job address:

During March or April, did you or NAME have a job that involved living or staying someplace else?

(If yes) **Who stayed someplace else?**

Did you stay at one place or more than one place while working?

(If one place) **What is the address where you stayed/stayed the most?**

Seasonal address:

Do you or does NAME have a seasonal or second home?

(If yes) **Who does?**

What is the address of your other home?

Other place:

In the past year, was there any other place you, or NAME stayed often?

(If yes) **Who stayed often at another place?**

What is the address where you stayed?

MODULE G: GROUP QUARTERS

In this module, the interviewer collects information about Census Day residence in GQ. GQ are inquired about separately, for the most part, since census residence rules state that the person is to be counted in a GQ if they were there on Census Day. This module is person-based. The following questions are asked for each person on the roster before being asked for the next person on the roster (after the first person, the question text is simplified for remaining people).

Please look at List A. The Census Bureau does a special count of people in places that house groups of people, such as nursing homes, jails, and emergency shelters. Even if you did not live there, did you spend even one night in any of those types of places around April 1?

List A includes the following types of places (This is paraphrased)
Dormitory or residence hall
Sorority or fraternity house
Military barracks
Military ship
Nursing home
Independent or assisted living facility
Correctional facility
Group home
Emergency shelter
Residential school for people with disabilities
Psychiatric hospital
Other

(If yes) **What type of place is it?**
What is the name of that place?
What is the address of that place?
Were you staying there ON Thursday, April 1?

MODULE H & I: QUESTIONS TO VERIFY ALL ALTERNATE ADDRESSES ARE COLLECTED FOR EACH PERSON

The goals of Module H and I are to verify that the interviewer has collected the appropriate addresses for each person on the roster. If a person had no other addresses collected within the instrument, then module H is used to verify this person only stayed at the sample address in the past year. If a person had two or more addresses, then module I verifies those addresses. In both modules there is the ability to add an address. In module I, there is the ability to delete an address.

Module H:
Just to confirm, the following people stayed at only one address in the past year: you, (and NAMEs). Is that correct?

(If no) **Who had more than one address?**
What is the other address where you stayed?

Module I:
I have collected these addresses for you: List addresses
Is that correct?
(If no) **What is the other address where you stayed?**
Which address(es) have you not stayed or lived at?

If another address is collected, the instrument collects whether the person moved or cycled, the move date if applicable, and any notes. These questions are identical to the questions in Module C.

MODULE J: CENSUS DAY RESIDENCE DETERMINATION FOR THOSE WITH MORE THAN ONE ADDRESS

In this module, the instrument collects data to determine Census Day residence status for people on the roster for whom there is not enough information yet to determine Census Day residence status. The goal of the module is to confirm where they lived, according to census residence rules, on Census Day. The questions in this module are person-based, that is, we ask all the questions in this module for each qualifying person before asking the same set of questions for the next qualifying person. If there are no people on the roster who fit the qualifications for the module (for example, no one has more than one address listed) the questions in this module are skipped.

Around April 1, where were you living and sleeping most of the time?
(If needed) **Where did you stay Thursday, April 1?**

The following questions are asked of those who are not inmovers or outmovers, have only one alternate address collected, and indicated they spend most of the time at that alternate address. This allows us to best determine how much time they spend at each location.

Which of the following categories most accurately describes the amount of time you stayed at the other place: A few days a week, A few weeks each month, Months at a time, Some other period of time?

Below are the followup questions to the previous question. Only the appropriate question(s) would be asked.

- **During a typical week, did you spend more days weeks/months at SAMPLE ADDRESS or the other place?**
- **During March and April, did you spend more weeks at SAMPLE ADDRESS or at the other place?**
- **In the past year, did you spend more months at SAMPLE ADDRESS or the other place?**
- **Were you staying at SAMPLE ADDRESS or the other place on April 1, 2010?**
- **Please tell me how much time you spent at each address in the past year.**
  Example of notes:
  - Only in Denver for a short visit during April; most of the time in Fayetteville
  - M-F in L.A. during March & April for job; rest of time in Stockton

MODULE K: INTERVIEW DAY RESIDENCE DETERMINATION

This module is similar to Module J in that the questions are only asked about people on the roster who fit into certain situations. The goal of this module is to make sure we have accurate interview day residence information. For these people, we ask questions to confirm where they live now (on interview day). The following questions are used to collect the interview residence information:

Currently, where are you living and sleeping most of the time?

If a new address is offered, the instrument collects the new address information using the standard questions in Module B and the type of place question (housing unit or GQ) in Module D. Additionally, if they say an alternate address, the instrument collects the length of stay at that address and any notes:

- **How long are you staying there?** (Response choices include: Less than a month, one – two months, three – four months, four – six months, more than 6 months, and not returning to SAMPLE ADDRESS)
- **Please tell me how much time you spent at each address in the past year.**

Example of notes:

- Only in Denver for a short visit during April; most of the time in Fayetteville
- M-F in L.A. during March & April for job; rest of time in Stockton

MODULE L: INMOVER ADDRESS

In this module, the interviewer collects more information about the inmover address (if there was one). The goals of the module are to collect enough information so that Census Bureau analysts can find the address in our records and for estimation purposes. If there was no inmover address collected within the instrument, this module is skipped. The section will be repeated if there is more than one inmover address (for example, two roommates moved into a sample unit from different places). If the address is a GQ, outside the country or we do not know the city and state, then these questions are skipped for that address.

Earlier you told me that you and NAME lived at INMOVER ADDRESS. I'm going to ask a few more questions about that address. Was there anyone else who lived there on April 1?

(If needed) **What are their names and approximate ages?**
Are any of these people related – NEW NAMES and you?

On April 1, 2010 was INMOVER ADDRESS owned with a mortgage or loan (including home equity loans), owned free and clear, rented, or occupied without having to pay rent?

What are the names of neighbors who lived near that place?

MODULE M & W: THE SAMPLE ADDRESS ON CENSUS DAY

In these modules, the interviewer collects information on housing unit status of the sample address on Census Day, and, if applicable, collects information on the residents of the unit on Census Day. Questions in this module are asked if the current residents did not live at the unit on April 1 or if it is not occupied now.

On April 1, 2010, was SAMPLE ADDRESS vacant or was it occupied?
The housing unit could be vacant, occupied, or not a housing unit.

(If the sample address is vacant or not a housing unit) **Which category best describes this unit? (**If vacant: **for rent, for sale, rent but not one lives there yet, sold but no one lives there yet, vacation home, migrant workers, other.**
If not a housing unit: **demolished, burned out, business, other, trailer moved from non-park location, empty trailer site, group quarter**)

(If occupied) **I need to ask some questions about the people who lived at SAMPLE ADDRESS on April 1 – things like names and ages. Do you know that information?**
(If yes) **What are the names of the people who lived here on April 1, 2010?**
I am going to show you the list of people I have recorded living here on April 1, 2010. Have I spelled all names correctly?

Then additional information is collected about the place they moved to and the date of the move (as in Module D), and the demographics (as in Module E) and the same questions in Modules F, G, I, and K to gather data on this roster.

MODULE N: RESPONDENT QUESTIONS

This module is the last module and contains questions for both the respondent and interviewer to wrap up the interview, or set a callback.

1. The type of respondent (landlord, neighbor, relative/caregiver, current occupant, former occupant, enumerator observation, other

2. Respondent information

What is your name?
In case we need to contact you again, may I please have your telephone number?
(In case we need to contact you again,) what is the best time to reach you?
(If proxy) **What is your address?**

3. If the interview requires a return visit, specific contact name, appointment date, time and proxy address if necessary is collected.

4. **This concludes our interview. Thank you Mr./Ms. LAST NAME very much for your cooperation.**

Attachment B: Person Interview Data Problem Log

#	Case/ Per/ Addr Lvl	Key variables/ Issue	# of Cases or People Affected	Description	Steps taken during production
1	Case	▪ Wrong OUTCOME[1] (OUTCOME= 203/209, but should be 326). ▪ Any records on 8502-8517 are off path data. ▪ Working as spec'd, but CURSTAT really should be 5 (instead of CURSTAT=1/4).	39	MOBILE VACANT. Cases coming through with a 203s/209s that are really seasonally vacant. They follow the main path in the front module. Upon reaching the HUTYPE question, they indicate they live in more transient location (other). We go to usual residence to confirm everyone living here has another place to live. If they say yes, then we end the interview and should at that point mark it as vacant. However, the instrument was not specified to reset the current status to vacant, and therefore the outcome is not set to vacant. So, the case looks like it is occupied without people.	None. Outcome is not used in processing.
2a	Case & Person	▪ Wrong OUTCOME (OUTCOME=201/203/208/209, but should be 326/327/333/334). ▪ May have Off path people on RT8502. [2] ▪ Wrong CURSTAT for **some** (CURSTAT 1/4, should be 2,3/5)	68	BAD VACANT. Vacant or seasonally vacant cases with OUTCOME of 201/203/208/209. May or may not have off path people listed on RT8502. Any people listed on 8502 are off path and the outcome is wrong.	None. See 2b.
2b	Case & Person	▪ Wrong OUTCOME (OUTCOME=201/203/208/209, but should be 326/327/333/334). ▪ May have off path people on RT8502. ▪ CURSTAT is blank.	51	BAD VACANT. This is a subset of the BAD VACANT cases in 2a for which the CURSTAT is blank. This subset issue is listed separately here because the blank CURSTAT may cause problems if the case is sent out for RI.	See OPS LOG 9/15; a new script was pushed to RI that updates the value of ORIG_CURSTAT with a value of 1 if it is empty so that the RI cases can transmit. Also, OPS LOG 9/14 provided special matching instructions for RI cases identified as being incorrectly coded as occupied in PI, but that we believe are vacant or NHU. Data were reviewed daily and FLD was notified immediately if any such cases were found.
3a	Case	Missing data on RT8502 - RT8531. Wrong OUTCOME (OUTCOME=203 but should be Type A Noninterview).	68	OCC_AVAIL ISSUE. Case has final outcome of 203, but only collected data in front and back. CAUSE: The instrument got in a loop at the OCC_AVAIL screen (front) and on the 2nd pass, the instrument did not jump back to the INTRO_S1 screen again as designed, but instead "dropped" into the back (R_Type). Since the case did not get to the back via the early exit function, the case was assigned as an outcome code of 203.	Data were reviewed daily and cases identified were sent to FLD for rework. 68 cases were sent back to be reworked. 12 left as is due to end of survey timing.

[1] See Section 5.4 for a list of all outcomes.

[2] RT = Record Type – Data was output in a ASCII file that was record typed by the type of information/section that the data was collected in.

#	Case/ Per/ Addr Lvl	Key variables/ Issue	# of Cases or People Affected	Description	Steps taken during production
3b	Case	Missing data on RT8502 - RT8531. Wrong OUTCOME (OUTCOME=203 but should be Type A Noninterview).	5	OCC_AVAIL REFUSAL. This is a variation of the OCC_AVAIL ISSUE in which the case went through OCC_AVAIL loop twice and entered REF at OCC_AVAIL on the 2nd pass. The instrument is not skipping to CUR_OCC as expected, but is instead skipping directly to the back (R_TYPE). Since the case did not get to the back via the early exit key, the case was assigned as an outcome code of 203.	Data were reviewed daily and cases identified were sent to FLD for rework. 5 cases to be reworked. None as part of final workload.
3c	Case & Person	▪ Wrong OUTCOME (OUTCOME=203/209 but should be Type A Noninterview). ▪ Missing data on RT8502 - RT8531.	20 found. 5 resolved in FLD.	MISCODED PARTIAL. Case appears to be a sufficient partial, but in actuality, the interview was terminated early and not all data were collected. Could be one of many issues. One known issue occurs when FR went to FAQ and also attempted an early exit or they attempted an early exit when already performing an early exit more than once. The early exit flag was not set, and therefore case gets a SUFFICIENT PARTIAL outcome instead of noninterview outcome code of 219.	Sent to rework depending on what the notes said (if according to notes, the respondent was not very hostile and rework could be expected to result in completion).
4a	Case	START_A is blank Outcome=21* (Note: COUNT_ATTEMPTS>0 for all)	3	BLANK START_A. Transmitted case that may not have actually been worked. Happens if the interviewer exits immediately at START_A; in that case all other data will be blank. One case has some data: the FR never reached a hh respondent (after 7 attempts) and eventually exited from START_A *after* deleting the value out of START_A. Hence, START_A is blank.	None. Start_A is not a key field and not used in processing.
4b	Case	COUNT_ATTEMPTS	5	COMPLETE CASES WITH ZERO ATTEMPTS. Case has a complete (20*) outcome and DOES appear to have complete data, but COUNT_ATTEMPTS = 0. All have ATTEMPT_TYPE filled. We believe this means that other count attempts for complete may be wrong. This problem and all problems in 4 are most likely due to reassigning the case as new.	None. Only needed for assessment.
4c	Case	COUNT_ATTEMPTS	323	NONINTERVIEW CASES WITH ZERO ATTEMPTS. Case has a noninterview (21*) outcome and COUNT_ATTEMPTS = 0. Most will also have ATTEMPT_TYPE=blank.	None. Only needed for assessment
4d	Case	ATTEMPT_TYPE	9	ATTEMPT_TYPE is blank. Transmitted case that may not have actually been worked. All have noninterview outcome. About half have START_A=2 and TRANSMIT=1, COUNT_ATTEMPTS=1.	None. Can use Respondent Type to determine.
5a	Person	Off path people on 8502 (OUTCOME is correct =Vacant/NHU)	160	VACANT WITH PEOPLE. Off path people mistakenly listed for a vacant/nhu unit. Assuming started listing the proxy's hh then went back and correctly identified status of actual sample address.	None. Went to Person Matching should be Nonmatch. Told Estimation for them to properly process.

#	Case/ Per/ Addr Lvl	Key variables/ Issue	# of Cases or People Affected	Description	Steps taken during production
5b	Case	Seasonally Vacant Error in Specification (Should have had WHO collected)	1182	Seasonally Vacant should not collect regular people but should collect WHO[3] information. Unfortunately one path in the front of the instrument did not send the cases to the WHO section. This is a specification issue. All other seasonal vacants collect WHO. All other flags are correctly set.	None. No correction possible after production began.
5c	Case	Seasonally Vacant got WHO (actually a good thing)	10	See 5b for mistake in spec but for some reason these cases mistakenly collected WHO. Unfortunately since does not meet proper WHO flow will be dropped in Post Processing like those listed in 5d	None. Dropped during Post Processiong
5d	Person & Address	Off path data on WHO RTs (8520, 8521, 8522, 8523, 8524, 8525, 8526, 8527, 8528, 8529, 8530, 8531)	169	OFF PATH DATA ON WHO RTs. Cases that were initially entered as vacant/NHU, entered in some WHO data before backing up and changing answers in Front to indicate that the unit is occupied and continuing down the main path.	N/a - this off path data will be removed during post processing.
6	Case	Missing Data at HU level	3	VACANT MISSING INITIAL INDICATOR INFO - Possibly interviewers entered FOUND_PROXY=2 (occupants away), collected people, then went back and changed FOUND_PROXY to 3 (vacant). Not sure why they did not skip to HU_STATUS1 at that point. So, have VACANT outcome, but don't meet "official VACANT/Not a Housing Unit" criteria, and DO have roster people. CURSTAT does seem to be right and we think outcome is right, but we do not know how. Side note these all have people so they are also counted in 5a	None.
7a	Person	Person Number missing on RT8503 (record present, but blank PERS[4])	7	OFF PATH DATA ON 8503 (Pers blank). Person number is blank on RT8503 (Roster Address), but is filled on 8502 (Person Records). Note that the # of records is good. Interviewer did a lot of changes to the roster and the roster alt addr in the section instead of just cleanly deleting them from the Roster Review Page. This created an error in the instrument because it created an address line to link to that person but they removed the person from the main list after giving the address and then moving forward. So, we have the person as a delete, but the roster address initially connected to them came out with no person number because when it moved forward that person did not exist anymore. The address data also do not make sense since they deleted most of the data out but not all of it.	N/A – Due to the bad linkage post processing drops this address which would have been done anyway since the person is a delete.
7b	Address	Off path data on RT8503 (PERS filled, but unclear why they went there)	4 cases, 14 people	OFF PATH DATA ON 8503. Complete data on 8503 (PERS, ADDR_ID, ROSTER_ADDR1, etc), but unclear why they went there. Names are complete and DELETE ne 1.	None.

[3] WHO = Whole Household Outmovers – This is section is used to collect Census Day household information when the interview day household was not there on Census Day. See Section 5.17.5.

[4] PERS = Variable Name for the Person Number assigned to each person listed to uniquely identify the person and any data associated with that person.

#	Case/ Per/ Addr Lvl	Key variables/ Issue	# of Cases or People Affected	Description	Steps taken during production
7c	Address	Missing RT 8503	6 cases, 11 people	The data indicate that it should have had an Roster Address but the data do not have one. Sometimes DELETE = 1 or ROSTER_REV = ' ' (i.e. a breakoff) which should make the address drop but not always. PP does not pick up people with Delete = 1	None.
8a	Case/Person	Person count is **greater** that the number of person records	13 cases	IMPROPERLY DELETED PEOPLE. Entering a person on the roster, sometimes entering address data as well (on 8503 or 8505), and going back later (but before the ROSTER_REV or REVIEW_LIST screens) to enter "999" in the FNAME line for that same person to delete the line rather than deleting the line at the ROSTER_REV/ REVIEW_LIST screen. 2 of them involve outmovers so there is an outmover address record but no person record.	None.
8b	Case/Person	Person count is **less** than the number of person records (includes missing)	168 cases	Percount is missing when problem 5a occurs. Percount is also missing for Problem 1a and 2a and 3a/b. Those are excluded here. BREAKOFF/MATCHING DK NAME - These are either because the person broke off before finishing the roster (and no percount was created) or at least one name was blank (entered 999 over a name originally entered or entered a DK name through more than one name so the instrument counted them as the same name).	None.
9a	Person	Off path outmovers on RT8502 [NOTE: it may look like these people are missing data on 8504-8514]	35 cases, 41 people	OFF PATH OUTMOVERS ON 8502. Added outmovers through outmover path, but went back and changed response to OUT_MOV = 2, 8, or 9. The RT8505 for that person is not there. Intention may have to delete people as well.	None.
9b	Address	Off path data on RT8505 (Present on 8505, missing on 8502)	14 cases, 18 people	OFF PATH DATA ON 8505. Added outmovers and started completing OUTMOV data, but later deleted these people by entering "999" in FNAME rather than deleting them from the REVIEW_LIST screen. Result is a record on 8505 for which there is no corresponding outmover on 8502.	None.
9c	Address	Outmovers missing demographics and data from RT8514	11 cases , 11people	MISSING OUTMOVER DEMOGRAPHICS & DATA ON 8514. Outmovers who are not asked the Review question(s) and are missing demographics. The respondent initially indicated they did not know about the outmover addresses, then went back much later in the interview and changed their response to indicate they DID know the address information. (The instrument used to allow such corrections but it resulted in inconsistent output, so the current instrument does not allow this type of late changes.) The current instrument output has the outmover names and outmover address information, but is missing demographics and was not asked the REVIEW series. (For some they do have demographics.)	None.

#	Case/ Per/ Addr Lvl	Key variables/ Issue	# of Cases or People Affected	Description	Steps taken during production
9d	Address	Outmovers with extra data (8505--8516)	3 cases ,3 people	OUTMOVERS WITH EXTRA DATA. The respondent initially indicated that they WERE knowledgeable about the Outmover, some data were collected, and later the interviewer went back and changed the response to say that they WERE NOT knowledgeable. If the INT made it past REVIEW_LIST before making the change, then the data are retained and the Outmover is treated as if NOKNOW=0 from that point forward (though in the output NOKNOW will = 1).	None.
10	Person	Off path data on RT8512 (Blank PERS)	173 cases, 180 People	OFF PATH DATA ON 8512 (blank PERS). Records with missing person numbers (PERS) on GQ RT (8512). ADDR_ID is filled for most. Similar to #7a (missing person number on 8503). They are saying Yes (1) at GQ_PLACE and entering some data, then going back and changing the answer at GQ_PLACE to No (2).	None.
11a	Case	Wrong ATTEMPT_TYPE (OUTCOME is correct - Vacant)	1054 cases	If a house is initially listed as seasonally occupied and then switched to seasonally vacant upon follow up questions, curstat is being updated (the important thing) but attempt_type is not switched to proxy. So Attempt_type indicates they confirmed at the sample address. For a very small number of cases the Attempt Type is off for curstat = 2 or 3 as well. (7 cases)	None.
11b	Case	R_TYPE (Respondent Type) values	639 Cases	The Value of Current Occupant (4) also is an option off the pick list on R_Type when talking to a proxy and can be picked even if should be talking to a proxy. (Spec /Interviewer Error Issue)	None
11c	Case	R_TYPE values	49 cases	R_Type indicates talking to a proxy when Attempt _type does not. Can ignore Noninterviews may not have gone through path again all the way to correct and Type B C since the AT would be wrong. *** NOTE: ALL but 5 of these fall into one of the problems 1 through 4c.	None.
11d	Case	Wrong CURSTAT (Current HU Status) value (OUTCOME and ATTEMPT_TYPE appear to be correct)	1 case	Nothing appears to be wrong, but CURSTAT=2 does not match A_T=1,3. This case does NOT appear to be a vacant case as with those in 11a. Need to research. NOTE: RESP_KNOWL=3.	None.
12	People	RELATION_REV (Relationship Review)	359 people	We are supposed to review relationship if the Reference Person is older than the person they indicate is a parent. Age is stored as 999 when DK. This is failing the check (999 is > any age) and relation review is being asked. This was an error in one of the early versions of the instrument and was fixed and tested. Do not know why it is back.	Discussed correction – Left as is - question appears just as confirmation.

#	Case/ Per/ Addr Lvl	Key variables/ Issue	# of Cases or People Affected	Description	Steps taken during production
13a	People	Mover = 2 (should be 5/6/blank)[5]	6 people	2 People have Mover = 2 and should be either 5 or 6 based on INMVR_DATE1. Others MOVER = 2 but HERE_CD = blank. These are breakoff people. No Mover should be set.	None.
13b	People	Mover = 3 (should be 2/1)	14 people	Mover was set to 3 if born on Census Day. Should have only been set to 3 for after CD.	None.
13c	People	MOVER=4 (should be 1/2)	28people	Person with ROSFLG of 2(Main Roster), HERE_CD = 1, no inmvr or outmover dates, addr_count zero, move = blank we think mover should be 1 – inst has it as 4 (not sure why). One thing of note all the cases have the same number of outmovers as those misclassified within the case. May be some part of an outmover person was checked in the instrument for that person that should not have been.	None.
13d	People	MOVER=1/4 (should be blank)	11 people	Some people that are missing the key data needed to set mover flag (MISCODED PARTIALS and Noninterviews) still have a mover value of 1 (for Main Roster People) or 4 (for Outmovers). Assuming that this is because 1/4 are set initially when these people go through the ROSTER_REV or REVIEW_LIST screens and only get updated as data are added. Values do not get blanked out if data are not added because the interview is terminated early.	None.
13e	people	MOVER=6 (should be 1/2)	10 people	Person should be 1 or 2 based on inmover. No reason to be 6 again all cases have an outmover and could be checking wrong person data for information.	None.
14	Case, Per, Addr	THROUGHOUT: All CHARACTER fields	None all cleaned in CMOCS or in PP	SPANISH & SPECIAL CHARACTERS. Although we were under the impression that we would not be receiving them, we found Spanish characters in the PI output. Currently, some cases are getting through the systems, but those that have Spanish characters in the address or contact information are failing check-in in CMOCs. This was also problematic because several software systems, specifically GEO's geocoding and the PeRMaRCS software systems, are not able to handle the special characters which would impact an already tight schedule for the CCM program.	1) FRs were instructed not to use any Spanish characters or accent marks when entering data into the instrument via the Ops Log. 2) Requirements for replacing Spanish and Special characters in PI data with English characters were added to the post-processing specifications through a formal CR. This requirement ensured that useable data would be passed to GEO's geocoding and PeRMaRCS software systems.
15	Address	ADDR_ID on RT8516	241 people	For some people, NOW_ADDR1 = '31' and the ADDR_ID associated with the record is equal to '1' . However, for the majority of cases NOW_ADDR1 = 31 --and-- ADDR_ID = 31.	A new Addr_ID was calculated by DSSD PI Team and forwarded to MARCS for use in Matching/Geocoding.

[5] Mover – Flag indicating to instrument if person is possible mover. It is only used in the instrument for controlling the flow of the questions and not in final residence coding. Values are 1 = Nonmover; 2 = Inmover; 3 = Born since Census Day; 4 = Outmover; 5 = Unresolved; 6 = Non-Resident.

#	Case/ Per/ Addr Lvl	Key variables/ Issue	# of Cases or People Affected	Description	Steps taken during production
16	Address	ADDR_ID on RT8515	724 people	If person says '32' at CD_ADDR1 and selects an address from the list at CD_ADDRESS2, the ADDR_ID gets filled with that address's ID. HOWEVER, if the person selects 31 from the list at CD_ADDRESS2, no ID gets assigned and should get an addr_id = 31.	A new Addr_ID was calculated by DSSD PI Team and forwarded to MARCS for use in Matching/Geocoding.
17	Case	Whole cluster of Type As	1 Cluster, 38 cases	Whole cluster of Noninterviews: We were denied access by the reservation in this cluster.	None.
18	Address	ADDRESS FIELDS (same addr_id, different address fields)	At least 252	We have found "irregular movements" in the address section that are leading to inconsistencies in the alternate address data. FRs are entering addresses, entering past the ADDR8/PR_ADDR13 field (where the code updates the external address file), and then mousing back to edit address components. When finished, they are mousing back to the last fields and therefore bypassing the key ADDR8/PR_ADDR13 field. As a result: • We get an "Updated" version of the address for the module in which the address was initially added and updated. • From that point forward through the GQ section, the interviewer will see the "Initial" version of the alternate address. If selected from the list, the "Initial" version of the alternate address will be output to the corresponding RT. Note that Interviewers CAN hit '0' and enter a new address at this point if they notice that the "Updated" version of the address is not in the list. • Interviewers will see both addresses listed on the REVIEW, CD, and NOW screens. The "Updated" version will be displayed first and the "Initial" version after that. These cases are not always easily identified and the count of addresses that had this problem may not account for all occurrences.	Created new address IDs through comparing the addresses and using number they chose on the pick lists for addresses. These address IDs were forwarded to Person Matching in time for the Matching Process.
19	Case	Two Clusters with High Noninterview Rates	2 Clusters. 137 cases out of 302.	An indian reservation only allowed access for one evening. The cluster is large and not completely contained in the reservation though and balances the noninterviews. The other cluster contains a large multi unit in NYC that the manager did not allow access.	None.
20	Case	STOPDATE	2 laptops	For two interviewers, their computers dates were set wrong. The dates are future dates.	None.
21	Case	INSTIME&TIMELENGTH_M	245 cases	245 cases have time (in minutes) over 200 going up to 983 minutes. This is unrealistic even with a high attempt count. This issue occurred in 2009 and is blamed on a interviewer closing a laptop without exiting and leaving the case open for hours or the instrument not properly shutting down due to many possible reasons.	None.

#	Case/ Per/ Addr Lvl	Key variables/ Issue	# of Cases or People Affected	Description	Steps taken during production
22	Case	Wrong ADDR_COUNT = -1, but should be 0	10 cases	NEGATIVE ADDR_COUNT. Addr_count = -1 for some cases. This is happening b/c the FR is deleting the address in REVIEW, and then going back to where they initially entered the address and changing the answer to say that they do not have the address. Therefore, the addr_count is changing from zero to minus one.	None.
23	Case	Sampled Group Quarters addresses not being reported as GQs in the PI	Unknown (At least 2 with one having a very high person count)	WORKING GQ CASE AS HOUSING UNIT. Found 2 instances where interviewer listed and worked what was a GQ case as if it was a regular HU case. 1) College owned condos (possibly multiple cases in cluster), 2) listed all people in nursing home (multiple people in one case).	None.
Source: Compilation of Daily Review of Data during Production and Team Review of Data					

Attachment C: Spanish Use

The following is the various distributions of use of Spanish in the PI instrument. As discussed in Section 5.15, we believe the use of Spanish may be slightly under reported and this is not a direct tracking of when they used the Spanish in the instrument. The states ranked in the top ten of their category are highlighted.

Table C- 1: Person Interview Distribution of Spanish Interviews by State Workload and Hispanic Origin Households (HH)

State	State Workload	Number of Spanish Interviews	Number of All Hispanic Households	Number of Cases that have any Hisp. (all or some)	Percent Span. Interviews for All Workload	All Workload Ordinal	Percent of Span. Interviews in Any Hisp. HH	Any Hispanic Ordinal	Percent of Span. Interviews in All Hisp HH	All Hispanic Ordinal
Alabama	2272	8	36	76	0.35	34	10.53	22	22.22	18
Alaska	1087	1	24	77	0.09	42	1.30	45	4.17	45
Arizona	6600	109	515	857	1.65	11	12.72	18	21.17	19
Arkansas	1394	6	34	61	0.43	29	9.84	23	17.65	24
California	19173	1493	4439	6039	7.79	1	24.72	6	33.63	7
Colorado	2463	13	215	383	0.53	27	3.39	40	6.05	43
Connecticut	1796	35	120	208	1.95	9	16.83	11	29.17	10
Delaware	1127	2	23	47	0.18	39	4.26	38	8.70	36
District of Columbia	1177	9	50	85	0.76	22	10.59	21	18.00	22
Florida	9560	584	1242	1620	6.11	3	36.05	2	47.02	2
Georgia	4720	41	178	267	0.87	18	15.36	13	23.03	15
Hawaii	4809	3	127	442	0.06	44	0.68	46	2.36	46
Idaho	1272	7	86	135	0.55	26	5.19	36	8.14	38
Illinois	6288	114	588	799	1.81	10	14.27	16	19.39	21
Indiana	3092	8	55	122	0.26	36	6.56	34	14.55	30
Iowa	1420	44	63	98	3.1	6	44.90	1	69.84	1
Kansas	1592	10	63	136	0.63	25	7.35	29	15.87	27
Kentucky	2069	0	23	43	0.00	47	0.00	51	0.00	51
Louisiana	2283	5	28	71	0.22	38	7.04	30	17.86	23
Maine	1170	0	5	24	0.00	48	0.00	48	0.00	48

State	State Workload	Number of Spanish Interviews	Number of All Hispanic Households	Number of Cases that have any Hisp. (all or some)	Percent Span. Interviews for All Workload	All Workload Ordinal	Percent of Span. Interviews in Any Hisp. HH	Any Hispanic Ordinal	Percent of Span. Interviews in All Hisp HH	All Hispanic Ordinal
Maryland	2964	12	80	176	0.40	31	6.82	33	15.00	29
Massachusetts	3252	25	203	285	0.77	20	8.77	25	12.32	34
Michigan	5259	14	90	203	0.27	35	6.90	32	15.56	28
Minnesota	3068	13	33	81	0.42	30	16.05	12	39.39	4
Mississippi	1653	18	41	67	1.09	14	26.87	4	43.90	3
Missouri	2839	3	38	83	0.11	41	3.61	39	7.89	39
Montana	1652	1	13	65	0.06	46	1.54	44	7.69	40
Nebraska	1184	3	31	67	0.25	37	4.48	37	9.68	35
Nevada	1508	95	244	332	6.30	2	28.61	3	38.93	5
New Hampshire	1067	7	31	55	0.66	24	12.73	17	22.58	16
New Jersey	4434	246	711	943	5.55	4	26.09	5	34.60	6
New Mexico	3756	47	572	860	1.25	13	5.47	35	8.22	37
New York	9572	250	991	1347	2.61	8	18.56	8	25.23	13
North Carolina	5091	45	200	302	0.88	17	14.90	14	22.50	17
North Dakota	1269	0	4	25	0.00	49	0.00	49	0.00	49
Ohio	5711	22	81	176	0.39	32	12.50	19	27.16	12
Oklahoma	1827	17	97	151	0.93	16	11.26	20	17.53	26
Oregon	1983	15	108	214	0.76	21	7.01	31	13.89	31
Pennsylvania	6166	23	173	285	0.37	33	8.07	27	13.29	32
Rhode Island	1155	33	110	153	2.86	7	21.57	7	30.00	8
South Carolina	2245	3	46	89	0.13	40	3.37	41	6.52	42
South Dakota	1472	1	15	38	0.07	43	2.63	42	6.67	41
Tennessee	3098	23	79	130	0.74	23	17.69	10	29.11	11
Texas	12371	612	2573	3401	4.95	5	17.99	9	23.79	14
Utah	1569	16	91	167	1.02	15	9.58	24	17.58	25
Vermont	1054	0	1	22	0.00	50	0.00	47	0.00	47

State	State Workload	Number of Spanish Interviews	Number of All Hispanic Households	Number of Cases that have any Hisp. (all or some)	Percent Span. Interviews for All Workload	All Workload Ordinal	Percent of Span. Interviews in Any Hisp. HH	Any Hispanic Ordinal	Percent of Span. Interviews in All Hisp HH	All Hispanic Ordinal
Virginia	4051	18	92	211	0.44	28	8.53	26	19.57	20
Washington	3855	54	181	373	1.40	12	14.48	15	29.83	9
West Virginia	1250	0	8	28	0.00	51	0.00	50	0.00	50
Wisconsin	3204	2	40	111	0.06	45	1.80	43	5.00	44
Wyoming	1375	11	84	145	0.80	19	7.59	28	13.10	33
Total	176318	4121	14975	22175	12.58	N/A	18.58	N/A	27.52	N/A
Source: Person Interview Final Version Output										

Table C- 2: Person Interview Distribution of Spanish Interviews by Regional Census Center and Hispanic Origin

RCC	RCC Workload	Number of Spanish Interviews	Number of All Hispanic Households	Number of Cases that have any Hisp. (all or some)	Percent of Span. Interviews for All Workload	All Workload Ordinal	Percent of Span. Interviews in Any Hisp. HH	Any Hispanic Ordinal	Percent of Spanish Interviews in All Hisp. HH	All Hispanic Ordinal
Boston (not PR)	13170	107	569	928	0.81%	9	11.53%	8	18.80%	10
New York	8350	411	1452	1880	4.92%	2	21.86%	3	28.31%	4
Philadelphia	13414	124	477	822	0.92%	7	15.09%	6	26.00%	5
Detroit	12220	36	179	407	0.29%	12	8.85%	12	20.11%	8
Chicago	12584	124	683	1032	0.99%	8	12.02%	7	18.16%	11
Kansas City	12140	93	328	610	0.77%	10	15.25%	5	28.35%	3
Seattle	15275	256	1305	2260	1.68%	5	11.33%	10	19.62%	9
Charlotte	16554	89	440	775	0.54%	11	11.48%	9	20.23%	7
Atlanta	16552	633	1456	1963	3.82%	4	32.25%	1	43.48%	1
Dallas	16307	635	2642	3539	3.89%	3	17.94%	4	24.03%	6
Denver	22861	297	1785	2943	1.30%	6	10.09%	11	16.64%	12
Los Angeles	16891	1316	3659	5016	7.79%	1	26.24%	2	35.97%	2
Total	**176318**	**4121**	**14975**	**22175**	**2.34%**		**18.58%**		**27.52%**	

Attachment D: List of Acronyms

Acronym	Definition
A.C.E.	Accuracy and Coverage Evaluation
CAPI	Computer Assisted Personal Interview
CCM	Census Coverage Measurement
CL	Crew Leader
CLA	Crew Leader Assistant
CMOCS	Coverage Measurement Operations Control System
CR	Change Request
CSM	Center for Survey Measurement
DAPPS	Decennial Applicant, Personnel and Payroll System
DMD	Decennial Management Division
DSSD	Decennial Statistical Studies Division
E Sample	Enumeration Sample
FDCA	Field Data Collection Automation
FOS	Field Operations Supervisor
GQ	Group Quarters
M&IE	Meals and Incidental Expenses
PES	Post-Enumeration Survey
PFU	Person Followup
PI	Person Interview
PR	Puerto Rico
P Sample	Population Sample
PSDB	Processing Systems Development Branch
RCC	Regional Census Center
RI	Reinterview
RT	Record Type
TMO	Technologies Management Office
WHO	Whole Household of Outmovers