

# Coverage Edit Followup

## **FINAL REPORT**

This evaluation reports the results of research and analysis undertaken by the U.S. Census Bureau. It is part of a broad program, the Census 2000 Testing, Experimentation, and Evaluation (TXE) Program, designed to assess Census 2000 and to inform 2010 Census planning. Findings from the Census 2000 TXE Program reports are integrated into topic reports that provide context and background for broader interpretation of results.

---

Dave Sheppard

Decennial Statistical  
Studies Division

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

*Helping You Make Informed Decisions*

---

Intentionally Blank

# CONTENTS

EXECUTIVE SUMMARY .....	vii
1. BACKGROUND .....	1
1.1 1990 Census .....	1
1.2 Census 2000 .....	2
1.3 What does this evaluation study? .....	5
2. METHODOLOGY .....	6
2.1 Census 2000 CEFU input files .....	6
2.2 CEFU evaluation files .....	6
2.3 Decennial Response File - Stage 2 .....	7
2.4 Hundred Percent Census Edited file with reinstated cases .....	7
2.5 SSD files containing respondent-reported data from the Internet Data Collection ...	7
2.6 Telephone interviewer debriefings .....	8
2.7 Applying quality assurance procedures .....	8
3. LIMITATIONS .....	8
3.1 Limitations of the input data .....	8
3.2 Limitations of the evaluation data .....	9
3.3 Limitations of the cost data .....	9
3.3 Other limitations .....	9
4. RESULTS .....	9
4.1 How many cases failed the coverage edit? .....	10
4.1.1 Over 2.5 million cases selected for CEFU .....	10
4.1.2 Fourteen files created, but only ten sent .....	11
4.1.3 Workload projections .....	12
4.1.4 Failure rates vary by form type .....	14
4.1.5 Failure rates vary by edit failure reason .....	15
4.1.6 Failure rates vary by language of form .....	15
4.1.7 Unsent cases not representative .....	17
4.2 How successful were we in contacting and completing cases? .....	19
4.2.1 Telephone number look up and telephone number quality .....	20
4.2.2 Making contact .....	21
4.2.3 Call backs needed .....	23
4.2.4 Completion of CEFU cases .....	25
4.2.5 Completion rate vs. Response rate .....	28
4.2.6 Completion rate differences by edit failure type .....	29
4.2.7 Completion rate differences by language of form .....	31

4.3	Were we successful improving coverage? .....	33
4.3.1	What was the net coverage gain by person? .....	33
4.3.2	Multiple adds, deletes and duplicates in one case .....	34
4.3.3	What happened to cases that were not completed? .....	34
4.4	What were other characteristics of the interview process? .....	38
4.4.1	Add probes .....	38
4.4.2	Delete probes .....	42
4.4.3	Names removed from the roster as duplicates .....	44
4.4.4	Name edits .....	46
4.4.5	Characteristics of persons added, deleted, or removed as duplicates .....	47
4.4.6	Characteristics of persons on the continuation roster of large household cases .....	48
4.4.7	Characteristics of households without a valid telephone number .....	51
4.4.8	Length of calls .....	51
4.4.9	Telephone Interviewer's use of CEFU instrument .....	52
4.4.10	Collecting person data .....	53
4.4.11	When were files returned to the Bureau from the contractor? .....	54
4.5	How costly were these coverage improvements? .....	54
5.	RECOMMENDATIONS .....	56
	References .....	57
	Appendix A: List of coverage edit followup probe questions .....	59
	Appendix B: Complete edit failure case deliveries .....	60
	Appendix C: More about cases with added persons and demographics of added persons ..	62
	Appendix D: More about cases with deleted persons and demographics of deleted persons	75
	Appendix E: More about cases with duplicated persons and demographics of duplicated persons .....	86
	Appendix F: More about cases with name edited persons .....	91
	Appendix G: Demographics of person one for added, deleted, or marked as duplicates ...	92
	Appendix H: Demographics of large household persons .....	95
	Appendix I: Demographics of persons cases with 'no valid telephone number .....	99
	Appendix J: Coverage questions from the 1990 and 2000 Census forms .....	102

## LIST OF TABLES

Table 1	Number of coverage edit followup cases selected by eligibility and whether they were sent to the contractor . . . . .	11
Table 2	Number of coverage edit followup cases sent in each delivery to the contractor compared to the projected file delivery sizes . . . . .	12
Table 3	Comparison of projected coverage edit followup workload to actual workload by edit failure type and form type . . . . .	13
Table 4	Coverage edit followup edit failure rate by form type and edit failure type . . .	15
Table 5	Coverage edit followup failure rate by edit failure reason . . . . .	15
Table 6	Coverage edit followup edit failure rate by language of mailback form . . . . .	16
Table 7	Coverage edit followup edit failure rate by language of mailback form and edit failure type . . . . .	17
Table 8	Likelihood coverage edit followup cases were not sent to the contractor by language of mailback form . . . . .	18
Table 9	Coverage edit followup edit failure rate for cases sent to the contractor compared to cases not sent by form type of mailback form . . . . .	18
Table 10	Results of the telephone appending operation by final disposition . . . . .	20
Table 11	Results of the telephone appending operation for cases with a final disposition of “no valid telephone number” . . . . .	21
Table 12	Distribution of calls per case to establish contact with the household for completed edit failure cases returned prior to July 1, 2000 . . . . .	22
Table 13	Distribution of calls per completed case after establishing contact with the household for cases returned prior to July 1, 2000 . . . . .	23
Table 14	Likelihood of callbacks for completed cases by type of edit failure . . . . .	24
Table 15	Likelihood of callbacks for completed cases by count discrepancy type . . . . .	24
Table 16	Likelihood of callbacks for incomplete cases by type of edit failure . . . . .	25
Table 17	Likelihood of callbacks for incomplete cases by count discrepancy type . . . . .	25
Table 18	Number and percent of the final disposition of coverage edit followup cases .	26

Table 19	Number and percent of coverage edit followup cases by disposition category by edit failure type .....	29
Table 20	Percent of coverage edit followup completed cases by disposition category and type of Census form submitted .....	30
Table 21	Number of coverage edit followup cases by disposition category and language of Census form. ....	31
Table 22	Percent of coverage edit followup cases by disposition category and language of Census form .....	32
Table 23	Distribution of person records in completed edit failure cases .....	34
Table 24	Distribution of completed cases by the number of person adds by completed edit failure case type .....	34
Table 25	Distribution of completed cases by the number of person deletes by completed edit failure case type .....	36
Table 26	Distribution of completed cases by the number of person duplicates by completed edit failure case type .....	37
Table 27	Percent of complete cases with adds by probe question and edit failure reason .....	39
Table 28	Number and overall percent of names added to the roster by coverage probe .	41
Table 29	Percent of complete cases with deletes by probe question and edit failure reason .....	42
Table 30	Number and overall percent of names deleted from the roster by coverage probe .....	44
Table 31	Percent of complete cases with duplicates by probe question and edit failure reason .....	45
Table 32	Number and percent of names deleted from the roster which represent the same person as another roster name by method of removal .....	45
Table 33	Distribution of cases with name edits for completed edit failure cases. ....	46
Table 34	Number of roster name corrections - names added, deleted, or removed as duplicates - by method of removal .....	46

Table 35 Percent of complete cases with at least one name edit by edit failure reason . . 52

Table 36 Central measures of interview length by edit failure type and form type . . . . . 54

Table 37 Cost summary for the Coverage Edit Followup Program and the Telephone  
Questionnaire Assistance Program . . . . . 55

---

Intentionally Blank



## **EXECUTIVE SUMMARY**

The Coverage Edit Followup operation for Census 2000 was used to increase within household coverage and improve data quality in two ways. First, it was used to collect person data for all persons beyond the first six in large households. Second, it resolved count discrepancies between the reported household population count and the actual number of data defined persons recorded on the census form.

The purpose of this evaluation is to document many aspects of the Coverage Edit Followup operation. It includes a look at the cases selected for the operation, documents our success at completing cases, profiles the persons we were able to collect data for during the operation, as well as explores the effectiveness of the operation's procedures and instrument.

### **How Many Cases Failed the Coverage Edit?**

While we had projected over 3.1 million coverage edit failure cases, we actually selected 2,544,072 coverage edit followup cases from Census 2000 mailback and Internet forms. Large household cases made up almost 55 percent of the coverage edit followup cases (1,395,623). The edit failure rate for large household cases was 1.7 percent.

Count discrepancy cases make up the rest (1,148,449). The edit failure rate for count discrepancy cases was 1.4 percent. Just over 60 percent (699,379 cases) of the count discrepancy cases were selected because the number of data defined persons on their form exceeded the respondent-reported household size. The rest of the count discrepancy cases were selected because the number of data defined persons on their form was less than the respondent-reported household size (449,070 cases).

### **How Successful Were We in Completing Cases?**

There were 1,251,971 cases completed during the Coverage Edit Followup operation. This was 53.5 percent of all the eligible and attempted cases. We were more successful completing large household cases (57.4 percent ) than count discrepancy cases (48.5 percent).

The largest reason for incomplete cases was our inability to contact the respondent by telephone. We made two attempts to obtain telephone numbers for cases where one was not correct or present on the mailback form. However, only 21.1 percent of the cases with changed telephone numbers were completed. Since there was no field followup, we were unable to complete any of the 562,049 cases where we could not obtain valid telephone numbers. This represented 24.0 percent of all eligible coverage edit followup cases.

### **Was the Coverage Edit Followup Instrument Effective?**

The coverage edit followup instrument was effective in its two main objectives: correcting incorrect rosters and collecting person data. However, some desired functionality was not available.

The coverage edit followup was not developed as a proposal for Census 2000 until after the dress rehearsal in an effort to address coverage concerns. Therefore, we did not have the opportunity to test and improve the operation by conducting it in a census-type environment prior to Census 2000. Many of the concerns that were raised at interviewer debriefings following the conclusion of the operation are worthy of consideration when planning similar operations in the future.

### **Were We Successful in Improving Coverage and Decreasing the Differential Undercount?**

The Coverage Edit Followup operation successfully improved coverage and decreased the differential undercount in Census 2000. In 232,777 cases, or 18.6 percent of all completed coverage edit cases, one or more persons were added, deleted, or removed as a duplicate. A total of 410,565 persons were added, deleted or marked as duplicates to correct the roster of a household.

The 152,683 persons who were added to the household roster during the operation were more likely to be members of traditionally undercounted populations than persons in the overall population enumerated in Census 2000. These persons were much more likely to be under 24, be of a race other than white, and to be Hispanic (especially Mexican). They were slightly more likely to be 65 years old or older, be male, and have the householder be an owner.

There were 257,882 persons who were deleted or removed as duplicates from the household roster during the operation. These persons were much more likely to be between 15 and 24 or over 65 years old and to be Black than persons in the overall population enumerated in Census 2000. They were slightly more likely to be of Hispanic origin, be female, and have the householder be an owner.

The Coverage Edit Followup operation actually resulted in a net loss of 105,199 persons compared to the originally completed Census 2000 Self Response forms. However, while the net improvement to the census from Coverage Edit Followup operation was a decrease in the population, it did improve the accuracy of Census 2000. Through the probing interview, the Coverage Edit Followup increased the likelihood that the 410,565 people who were added, deleted or marked as duplicates were counted in the correct household.

### **What are the Recommendations?**

Given the results and limitations of the data, here are some recommendations for the Coverage Edit Followup operation in Census 2010:

- ▶ Continue to conduct a coverage edit followup operation in future censuses. Include count discrepancy cases and large household cases, as well as other cases we can identify as having a significant possibility of coverage problems.

- ▶ Develop ways to increase the completion rate for Coverage Edit Followup operations in the future. We should:
  - Conduct a field followup for cases we do not reach by telephone. This is especially necessary for Puerto Rico and other areas which typically do not have telephones; all cases deserve a followup.
  - Improve our ability to obtain correct telephone numbers for the respondent.
  - Conduct a refusal conversion operation by telephone or field followup to improve the completion rate.
  - Allow interviewers to leave a message when respondents are unavailable so they may call us back to complete the followup.
  
- ▶ Improve case file creation, management, software testing and transmittal procedures of input and output files to avoid loss of data and to ensure information is available to conduct interviews as planned. We should:
  - Improve testing of the universe selection software to avoid selecting ineligible cases for followup and to avoid missing key variables on the input files.
  - Ensure that attempted cases are representative of the entire universe of coverage edit cases in the event the full originally selected universe cannot be followed up.
  - Improve testing and monitoring of files received from contractors in the future to ensure their completeness and accuracy.
  
- ▶ Improve the design of the coverage edit followup instrument to improve effectiveness and reduce respondent burden. We should:
  - Allow telephone interviewers' input into the design of the survey instrument earlier in the development process.
  - Tailor the probe questions to the specific edit failure reason based on the results of this operation during Census 2000 and other relevant research.
  
- ▶ Collect evaluation data in future census tests of coverage followup operations to help improve the methodology used to conduct followup interviews. Ensure we can:
  - Collect and analyze the number of call attempts for use in establishing contact with households as well as the number of attempts needed to complete cases in a telephone followup operation.
  - Collect and analyze program cost data to better understand the true cost of the coverage improvements gained from coverage edit followup.
  
- ▶ Assign the final household size for count discrepancy cases not completed during coverage edit followup by more closely mimicking the results for completed cases in Census 2000.

---

Intentionally Blank

# 1. BACKGROUND

## 1.1. Coverage edit followup in the 1990 Census

A Coverage Edit Followup (CEFU) operation was conducted as part of the 1990 Census. These coverage edits relied on a comparison of respondent supplied, office coded, and computer interpreted data.

The universe for this edit was all mail return and enumerator short and long forms. Two coverage questions were on each questionnaire. These questions can be found in Appendix J. A combination of clerical and computer edits of eligible cases were completed to identify questionnaires that met the criteria for the CEFU.

Questionnaires failed edit if any of the following occurred:

- The questionnaire was blank or had only housing questions answered (mail return only).
- The respondent had seven or more persons listed on the roster (mail return only).
- The respondent indicated that the household had a usual home elsewhere (WHUHE) as shown in Appendix J.
- There was a population count discrepancy between the number of person columns completed and the number of persons on the household roster.
- The respondent had problems deciding who should be included on the questionnaire (mail return only) determined by their responses to the two coverage questions H1a and H1b as shown in Appendix J.

Because the questionnaire only had room to enumerate seven persons, all mail-return questionnaires that had entries in all seven person columns failed edit as there may have been more persons yet to be counted.

Cases failing for any of the above reasons were considered 'coverage problems' and were marked for telephone followup. Cases were resolved by telephone followup interviewers following the instructions in the District Office Telephone Followup Manual using the Questionnaire Reference Book (QRB). The instructions explained how to resolve CEFU cases, but did not provide a script or series of questions to ask the respondent.

When respondents could not be reached by telephone, the cases from mail returns were referred to the District Offices for enumerator field visits. Enumerator returns not contacted by telephone were not sent to the field. Finally, the telephone and/or field enumerator used the respondent or enumerator completed questionnaire during the followup interview. All followup work was done by Census Bureau staff.

Due to budget constraints, no formal evaluation was done of the effectiveness of this operation after the 1990 census.

## 1.2. Coverage edit followup in Census 2000

A CEFU operation was conducted as part of Census 2000. This telephone operation was used to improve within household coverage and data quality in two ways. First, it was used to collect person data for all persons beyond the first six in large households (the maximum number of people we could collect data for on mail back forms in Census 2000 was six). Second, it resolved count discrepancies between the reported household population count and the actual number of data defined persons recorded on the census form. Prior to collecting person data, a series of probes were asked for all CEFU cases. These probes were designed around the residence rules and allowed the respondent to identify persons that should be added to or deleted from the household roster as reported on their census mail back form. This would then more accurately represent the actual household composition.

The universe for this edit consisted of all mail return short and long forms (SF and LF) as well as certain Be Counted forms (BCF) and Internet data collection (IDC) responses processed by June 8, 2000. Census 2000 forms of these types had several language versions that were eligible for CEFU. In addition to the standard English form, there were forms in Spanish, Chinese, Korean, Tagalog, and Vietnamese. The forms in both Spanish and English used in Puerto Rico were also eligible. A computer edit of these cases was done to identify eligible questionnaires that met the criteria for the CEFU.

Enumerator forms, used for nonresponse followup, coverage improvement followup, and the update/enumerate operation, were not eligible for CEFU because it was unnecessary. When enumerator forms were used, information was collected for household members in large households using continuation forms. Also, any enumerator forms which had a count discrepancy should have been screened out by the crew leader and returned to the field for rework. There were also coverage questions on enumerator questionnaires, which were not on the forms eligible for coverage edits, to help ensure the household roster was correct.

These coverage edits relied on comparisons of respondent supplied and computer interpreted data. The Census 2000 coverage edit failures were determined using the respondent-reported household size, the number of data defined persons on the roster, and the number of continuation roster names. Persons were determined to be data defined during previous Census processing based on the number of data items supplied for that person. There were two types of coverage edit failures: count discrepancy followup cases (CDFU) and large household followup cases (LHHFU). There were two CDFU reasons:

- **Count Discrepancy - High data defined persons (HDDP)** where there were more data defined persons than the reported household size (for SF, LF, BCF, and IDC) on the form. For example, if the household size was listed as four by the respondent, but six persons were data defined on the form.
- **Count Discrepancy - Low data defined persons (LDDP)** where there were fewer data defined persons than the reported household size (for SF, LF, BCF, and IDC) on the form. For example, if the household size was listed as three by the respondent, but only two persons were data defined on the form.

There were two LHHFU reasons:

- **Large Households (LHH)** for SF, LF and IDC forms where the reported household size or the sum of data defined persons and continuation roster persons was greater than six. The BCFs failed as large household cases if the reported household size or the sum of data defined persons and continuation roster persons was greater than five.
- **Possible Large Households (PLHH)** for SF, LF, and IDC forms with exactly six people listed but the total person count on the form was left blank.

### **Conducting Coverage Edit Followup Interviews for Census 2000**

The Census Bureau staff specified instrument requirements and selected the cases for CEFU from the universe of eligible cases. However, the actual followup of these cases was contracted to Electronic Data Systems (EDS). The EDS assembled the resources to conduct the entire telephone followup operation. Its role included:

- creating a computer assisted telephone interview (CATI) instrument
- reserving and monitoring the work of multiple call centers
- obtaining and training telephone interviewers
- creating and controlling the infrastructure to control the flow of data from receiving input files to returning the completed cases to the Census Bureau.

The CEFU attempted to contact all households by telephone that failed edit. Telephone interviewers, also known as agents, used a browser-based desktop application. The instrument included a series of help sources called the knowledge database. There was no field visit or enumerator followup for CEFU cases that were not resolved over the telephone.

In contrast to the CEFU operation in the 1990 census, the CEFU operation was very scripted in Census 2000. Questions were to be asked verbatim to assure consistency from interview to interview, especially since interviewing occurred at thirteen different call centers. In addition, the telephone interviewer did not have the respondent completed questionnaire; instead, they only had the relevant data from the questionnaire.

The interviewing procedure began when the auto dialer system attempted to contact a household in the CEFU universe. If the telephone was not answered, the case was recycled for additional calls at a later date. If a household was reached, the telephone interviewer determined whether the correct household was reached and if so, whether an eligible respondent was available and able to conduct the interview at that time.

According to our requirements, only persons listed as person one or person two on the household roster of the mail back form were eligible to respond to the CEFU interview. This was done to increase the likelihood that the respondent would be knowledgeable enough about the household to provide correct responses. If an eligible respondent was available, the interview was conducted. If not, the case was recycled for additional calls at a later date.

The telephone interviewer would read the respondent-reported household roster to the eligible respondent. The telephone interviewer then asked a series of nine questions designed to ensure that the household roster was complete and correct (see Appendix A). The first five of these probes were based on the Census 2000 residence rules and designed to determine if additional persons should be added to the household roster. The last four probes similarly were designed to determine if persons on the household roster should not be listed according to the Census 2000 residence rules.

For each of these nine coverage probe questions, a similar flow of questions was followed. For example, there were questions designed to add persons left off their mailback Census 2000 form in error. After being read the household roster, the respondent was asked if a person with particular characteristics (child, roommate, and so forth...) was living or staying there around the beginning of April and was not included on that roster. If so, we then asked for that person's name. If a name was offered, we then confirmed with the respondent that this person was living or staying there most of the time as of April 1. This multi-stage approach allowed the respondent to consider more possible residents while we defined the criteria within our followup questions.

In addition, the respondent could interrupt the interview at any point to make corrections to the household roster. Telephone interviewers would then take the appropriate action using the interrupt options. There were four interrupt options: adding a name to the roster, deleting a name from the roster, indicating that more than one roster name represents a particular household member, and editing the name of a person on the household roster. Upon the completion of this action, the interview was resumed where it was left off.

Once all the probes were asked and answered, the case was considered count complete because we had confidence that the number of persons on the household roster was correct. If data needed to be collected for one or more of the persons on the household roster, they were collected after the nine probes were asked. If a person on the roster was confirmed to be a delete or a duplicate, a flag was set and the person record was retained. Otherwise, the CEFU interview ended.

Due to delays in development and testing, the start of the program was delayed. Note that the planned finish was an arbitrary date since there were not any operational dependencies that dictated we finish by then. In fact, EDS was told from the start that this date was open for extension.

A contingency for a second phase of the CEFU operation was planned to allow a mechanism to potentially raise the overall completion rate. It was thought this could be achieved by contacting the non-interviews as well as improving the coverage of the non-English speaking population. The requirements for reallocating cases that need to be retried, ensuring the allocation of remaining cases, and closing out the operation were specified in advance. This contingency, referred to as phase two, was implemented between August 1, 2000 and August 12, 2000.



---

Planned start: April 5, 2000	Planned finish: June 19, 2000
------------------------------	-------------------------------

---

Actual start: May 8, 2000	Actual finish: August 13, 2000
---------------------------	--------------------------------

---

### 1.3. What does this evaluation study?

The overall objective of this evaluation is to look at several aspects of the CEFU operation for Census 2000. We look at the workload, completion rates, effectiveness of the CEFU instrument, coverage gains, and the cost. Additionally, we look at the demographics of several groups of household members who completed the CEFU interview. This included the people who were added, deleted, or removed as duplicates from the household roster during CEFU as well as those people who were on the continuation roster and had their demographic data collected during CEFU.

To get a more complete understanding of the planning, the issues, and the outcomes of the coverage edit followup operation for Census 2000, this report should be read in conjunction with the following three reports prepared by the Decennial Systems and Contracts Management Office (DSCMO), the Planning, Research, and Evaluation Division (PRED), and the Decennial Management Division (DMD):

Census 2000 DSCMO General Memorandum Series #01-01, dated June 12, 2001, from Michael J. Longini, Chief, DSCMO, to Distribution List, Subject: *Telephone Questionnaire Assistance (TQA) and Coverage Edit Follow-up (CEFU) Lessons Learned for Census 2000 - Revised*, DSCMO

U.S. Bureau of the Census, 2002d, *R.1b - Coverage Edit Followup System Requirements Study*, PRED, Census 2000 Evaluations

U.S. Bureau of the Census, 2002e, *Census 2000 Coverage Edit Follow-up Comprehensive Operational Assessment*, Final Draft April 15, 2002, DMD

The results of these four evaluations will aid planners for the 2010 census in designing coverage related operations.

## 2. METHODOLOGY

We used six data sources for this analysis:

- The Census 2000 CEFU input files,
- The CEFU Evaluation files,
- The Decennial Response File – Stage 2 (DRF2),
- The Hundred Percent Census Edited File with the reinstated cases (HCEF\_D'),
- Data files from Systems Support Division (SSD) containing respondent-reported data from the Internet data collection, and
- Telephone interviewer debriefing results.

Each file will be addressed as to how they were used within this report.

### 2.1. Census 2000 coverage edit followup input files

The CEFU input files were used to answer questions about the number and types of cases that failed the coverage edit. Fourteen files were created, for the most part one per week. These fourteen files were created by the DSCMO from March 23, 2000 through June 8, 2000. These files were created to send the CEFU cases to the contractor as input into the operation.

The CEFU cases were only selected from eligible cases data captured by June 8, 2000. However, only the first ten of these files were sent to EDS for interviewing (see Section 4.1.1 to find out why). A total of 2,506,998 cases were contained in these ten files while 92,486 cases were in the four files that were not sent.

Additionally, we determined that these fourteen files included 55,412 cases that were not eligible to be selected. These cases were ineligible because they did not include any name information for the first or second person listed on the mail back Census form, which was a requirement for the universe selection. The ten files we sent contained 48,109 ineligible cases while the four files that were not sent contained the remaining 7,303 ineligible cases.

For the purposes of this analysis, the ineligible cases are not considered. Therefore, the universe of cases appropriately selected for coverage edit followup is 2,544,072. Of those, 2,458,889 cases were sent to EDS and the 85,183 eligible cases from the final four files were not sent and never had a opportunity to be completed.

### 2.2. The coverage edit followup evaluation files

The EDS transmitted output files to us almost daily during the CEFU operation. These files served two purposes. The production files contained the census data from the completed CEFU interviews. The evaluation files were created in order to evaluate the CEFU operation.

Production data files were divided into short form (including Internet and BCFs) and long form cases. For every file transmission, we received one of each file provided there were both short form and long form cases completed for that delivery. Production files were NOT used to answer questions in this evaluation.

The CEFU evaluation files specified by the Decennial Statistical Studies Division (DSSD) were used to answer questions about most of the CEFU study plan questions. Evaluation files were divided into household level and person level data files. These files were created by EDS and sent to DSCMO on a daily basis. Each transmission included both household and person level files. We received the first evaluation files on June 2 and the last on August 16. We received 66 pairs of files containing completed cases during this time.

Once the operation had ended, we received two additional pairs of files. These files contained non-interview cases -- those with resolved status codes as well as those with interim codes. Had the operation continued beyond its end date, these cases with interim codes would still have been called in attempt to complete the CEFU interview.

We eliminated some of the returned cases from our analysis because of incomplete information. We eliminated 9,370 household level records because there were no corresponding person records returned to us on the evaluation files and 13,357 household level records because there were person data but no corresponding household records returned to us on the evaluation files. And, after matching the evaluation files with the input files originally sent to EDS, we found that 97,742 eligible cases which we had sent to EDS were never returned with any status on the evaluation files.

Therefore, 4.5 percent (120,469 of the 2,458,889) of CEFU cases were removed from this analysis. This resulted in 2,338,420 cases appropriately sent to EDS and returned to the Census Bureau with a complete, incomplete, or interim case disposition on the evaluation files.

### **2.3. Decennial Response File – Stage 2 (DRF2)**

A file of all CEFU evaluation cases was created and matched to the DRF2. Information appended from this file was used to determine which cases were submitted on Asian Language Census forms as well as to indicate the tenure status of each housing unit.

### **2.4. The Hundred Percent Census Edited File with the Reinstated Cases (HCEF\_D')**

A file of all CEFU evaluation cases was created and matched to the HCEF\_D' file. Information about how the respondent answered the tenure question was appended from this file.

### **2.5. Data Files from Systems Support Division containing respondent-reported data from the Internet Data Collection**

The System Support Division (SSD) Internet data files were used in conjunction with CEFU input and output files to determine which CEFU cases submitted their data through Internet Data Collection.

## **2.6. Telephone Interviewer Debriefing**

Several debriefings were held with people from all stages of the CEFU operation. Two debriefings involved telephone interviewers and their supervisors. One was held in Troy, Michigan on August 14, 2000 and another at the Census Bureau's National Processing Center in Jeffersonville, Indiana on November 9, 2000. The notes from these debriefings were used to answer questions about the CEFU instrument's effectiveness from the telephone interviewer's perspective.

## **2.7. Applying Quality Assurance Procedures**

We applied quality assurance procedures throughout the creation of this report. They encompassed how we determined evaluation methods, created specifications for project procedures and software, designed and reviewed computer systems, developed clerical and computer procedures, analyzed data, and prepared this report.

# **3. LIMITATIONS**

There were several limitations to the data used for this evaluation. There were inconsistencies between and among the evaluation and input files.

## **3.1. Limitations of the evaluation data**

Some data had to be eliminated from our analysis because there were no corresponding person and household level data. A total of 120,469 cases, as indicated earlier in Section 2.2, were removed from the analysis.

Some data we had specified to receive were not provided to us. For example, elapsed time of call, including the time spent during each call to a household, was never programmed successfully by the contractor.

Due to limitations of the existing system, we knew some evaluation data would be overwritten prior to the start of the second phase of the operation. Therefore, for cases that were active during phase two and completed, we do not know the cumulative number of call attempts needed to make contact or to complete the interview. Also, for cases that were active during phase two and not completed, we have no idea how many calls were made attempting to make contact and/or to complete the call.

Most of the persons added or deleted from household rosters occurred through the interrupt screen. No information was collected about the reasons for these actions, so we only know reasons for the adds and deletes from cases where it was the result of one of the nine coverage probes.

### **3.2. Limitations of the input files**

Some of the identification variables on the input files were missing. Information about language of mail back form and form type were not filled on the input files we gave to EDS, which created difficulties during the operation as well as during the evaluation. Alternative sources were found for this information and were appended to the input and evaluation files.

### **3.3. Limitations of the cost analysis**

Included in the contract for CEFU was the Telephone Questionnaire Assistance (TQA) program. This program was a short duration program implemented to assist the public in completing their census forms or obtaining information about the census. The requirements for conducting cost analysis of the CEFU for evaluations was specified after the award of the contract and the agreements on how to report costs for the TQA program. Therefore, some of the item costs for both the inbound (TQA) and outbound (CEFU) components were not billed separately by the contractor. We were not able to accurately report the separated costs for the CEFU program for these item costs. Moreover, we were not able to report the true value of the total cost of the CEFU operation. In addition, headquarter costs were not included in the cost figures.

### **3.4. Other limitations**

The demographic data for persons enumerated during CEFU, persons removed during CEFU, as well as householders without a valid telephone number in CEFU were based on unedited data. However, the data for persons in the overall Census 2000 population, used for comparison purposes in this report, were based on edited data. The assumption is that they are distributed like the cases with observed values. If not, they could distort the distribution.

## **4. RESULTS**

The CEFU operation for Census 2000 was a very complex operation. While only one instrument was used for all the cases, there were a wide variety of differences among these selected cases. Four form types - short forms, long forms, Internet forms, and Be Counted forms - were eligible for selection. Forms in six languages - English, Spanish, Chinese, Vietnamese, Korean, and Tagalog - were eligible for selection. Each eligible form could have failed the edit for one of four reasons - large household, possible large household, count discrepancy with high data defined person count, and count discrepancy with low data defined person count.

Complete CEFU cases either had the household roster remain the same or changed. If it was changed, there could be persons added to it or persons deleted from it, or both. Each name added or removed from the roster is linked to one of thirteen reasons, usually coverage probes questions, which led to the change. Data were collected for two types of people: persons listed on forms associated with large household cases as well as persons who were added during the CEFU operation.

The CEFU instrument had several aspects worth noting. Each of the nine coverage probes had a three step unfolding structure leading to roster changes. Especially since no dress rehearsal of these methods was conducted prior to Census 2000, we wanted to learn as much as we could to aid us in planning similar operations in the future.

This report will cover many aspects of the CEFU operation for Census 2000. There are seven subsections in the results section. First, Section 4.1 will describe the workload of the operation. Section 4.2 details how successful we were contacting and completing these cases. Coverage gains will be discussed in Section 4.3. Section 4.4 looks at some measures of the effectiveness of the CEFU instrument. Demographic characteristics of persons on large household continuation rosters for whom we collected data during CEFU are contained in Section 4.5. Section 4.6 looks at a few other characteristics of the CEFU interview process. Section 4.7 attempts to decipher the costs associated with the CEFU operation.

#### **4.1. How many cases failed the coverage edit for Census 2000?**

In this section, we present the workloads associated with the coverage edit followup operation for Census 2000. We will show how many cases were selected by DSCMO from the daily normalized files which were created from data capture records. As appropriate, counts and rates are shown by selection date, CEFU eligibility status, edit failure type, edit failure reason, type of Census form and language of Census form.

##### *4.1.1. Over 2.5 million cases selected for coverage edit followup*

As Census 2000 forms were processed at the data capture centers, the data were sent to DSCMO on a daily basis. Coverage edit followup cases were selected from four eligible form types - mailback short forms, mailback long forms, Be Counted forms, and Internet data collection submissions. Selections were made from all eligible forms processed by June 8, 2000. There were a total of 2,599,484 cases selected by that date.

As shown in Table 1, not all of the selected cases were in fact eligible or even necessarily used. At the start of the operation, EDS realized that we had sent them some cases that were ineligible according to our universe specifications. For example, our specifications excluded cases where there was no last name reported for person one. Without a last name for person one, we had no one to ask for when we called the household. The EDS screened out these cases before distributing them to the call centers and no attempts were made to contact these households. There were 55,412 cases that were selected and transmitted to EDS in error.

Additionally, a decision was made by Census management to stop sending selected CEFU cases to the contractor as of May 15, 2000. This was made for two reasons. First, it was believed we had already delivered more cases than EDS could handle prior to the planned end date of the operation. Second, management knew that all of these cases were late mail returns and would be included in the nonresponse followup universe. Therefore, 85,183 eligible cases were not sent and no attempts were made to contact these households through CEFU.

The remaining eligible cases, 2,458,889, were distributed by EDS to the thirteen call centers for interviewing.

**Table 1. Number of coverage edit followup cases selected by Eligibility and whether they were sent to the contractor**

	Coverage Edit Case		
	Total Selected	Sent	Not Sent
<b>Total Cases Selected</b>	<b>2,599,484</b>	<b>2,506,998</b>	<b>92,486</b>
Eligible Cases Selected	2,544,072	2,458,889	85,183
Ineligible Cases Selected	55,412	48,109	7,303

Source: CEFU input files - variables *sample* and *file*

#### 4.1.2. Fourteen files created; only ten sent

The first ten files, containing 2,506,998 eligible cases, were delivered to EDS on a mostly weekly basis from late April through mid May 2000. The final four files, containing 92,486 cases, were never sent. No attempts were ever made to interview these cases through CEFU.

Table 2 shows a comparison of the number of cases we selected compared to the number we expected to deliver over the time of the selection processing. This table includes all cases sent to the contractor, including cases later determined to be ineligible.

We had been overly optimistic about how many cases would be processed early. Since fewer cases were processed early we therefore had fewer CEFU cases selected early. However, the operation was delayed and did not start until May 8, 2000. On that date, our projections of how many cases would be selected by that date were actually very close to the actual numbers. We had projected 2,225,000 cases and had actually chosen 2,235,418 by May 8 - a difference of only 10,418. Overall, we selected and delivered 603,002 cases less than we had projected for the entire program.

**Table 2. Number of coverage edit followup cases sent in each delivery to the contractor compared to the projected file delivery sizes**

<b>Delivery Date</b>		<b>Number of Cases Selected</b>	<b>Projection</b>	<b>Week of</b>	<b>Difference</b>
March 23, 2000	sent	140,922			
March 28, 2000	sent	173,574			
April 5, 2000	sent	319,563	950,000	April 5, 2000	(315,916)
April 6, 2000	sent	25			
April 13, 2000	sent	534,959	350,000	April 12, 2000	184,959
April 20, 2000	sent	383,240	325,000	April 19, 2000	58,240
April 26, 2000	sent	6,065	325,000	April 26, 2000	(58,276)
April 27, 2000	sent	260,659			
May 5, 2000	sent	416,411	275,000	May 3, 2000	141,411
May 12, 2000	sent	271,580	275,000	May 10, 2000	(3,420)
May 19, 2000	not sent	55,157	275,000	May 17, 2000	(219,843)
May 26, 2000	not sent	20,763	275,000	May 24, 2000	(254,237)
June 2, 2000	not sent	9,768	50,000	May 30, 2000	(40,232)
June 8, 2000	not sent	6,798	10,000	June 7, 2000	(3,202)
	<b>sent</b>	<b>2,506,998</b>			
<b>Totals</b>	<b>not sent</b>	<b>92,486</b>	<b>3,110,000</b>		<b>(603,002)</b>

Source: CEFU input files

*4.1.3. Workload projections for large household cases were close; those for count discrepancies were not*

When the contract for the CEFU operation was first awarded, decisions about the universe had not yet been made. Without that information, the workload was projected to be between 580,000 and 4.5 million. Soon thereafter, a workload estimate of 3,110,000 was provided to the contractor spread over ten consecutive weeks. By six months before the planned start of the program, the universe had been defined as 3,250,000 cases delivered over nine consecutive weeks.

Table 3 shows the comparison of the overall projected workload to actual workloads for each form type and edit failure reason by week of planned delivery. Overall, 2,544,072 cases were selected for coverage edit followup, 565,928 less than the 3,110,000 case projection.

Our projections for large household cases were very good while the ‘guess’ we made about the count discrepancy workload was not so good. We based the projections for large household



cases on the household size data from the 1990 Census, estimates of the mail response rate, and assumed that one in six mailback forms would be a long form. However, our count discrepancy projections are another story. For the purposes of planning staffing levels, we made an assumption that there would be 1.8 million count discrepancy cases. While we ended up with many fewer count discrepancy cases overall, there were actually many more long forms count discrepancy cases than we had projected.

**Table 3. Comparison of projected coverage edit followup workload to actual workload by edit failure type and form type**

Type of Edit Failure by Form Type	Number of Coverage Edit cases	
	Selected	Projected
<b>Total of all Addresses</b>	<b>2,544,072</b>	<b>3,110,000</b>
<b>Large Household</b>	<b>1,395,623</b>	<b>1,320,000</b>
Short Forms	1,231,726	1,120,000
Long Forms	156,729	200,000
Be Counted Forms	5,941	*
Internet Forms	1,227	*
<b>Count Discrepancy</b>	<b>1,148,449</b>	<b>1,800,000</b>
Short Forms	790,470	1,530,000
Long Forms	357,369	270,000
Be Counted Forms	n/a	n/a
Internet Forms	610	*

\* Projections were not made using Be counted forms or Internet forms  
 - These are small workloads and were added to the operation after the initial workloads were projected

Source: CEFU input files

#### *4.1.4. Four form types with varying failure rates*

Table 4 shows the overall failure rate for each form type broken out by basic edit failure reason. There were 82,008,049 Census 2000 forms processed by June 8, 2000, that were eligible to be selected as CEFU cases. Of these, 2,544,072 forms or 3.1 percent, were correctly chosen as CEFU cases. There were more large household cases (1,395,623 or 1.7 percent of eligible cases) than count discrepancy cases (1,148,449 or 1.4 percent of eligible cases).

Eligible long form cases were two and a half times more likely to fail the count discrepancy edit than short forms. These differences may be partially explained by the different criteria for choosing edit failures for short and long forms and the fact that if a case met the criteria for both count discrepancy cases (CDFU) and large household followup cases (LHHFU), the case was listed as a LHHFU case.

For short form cases, a case failed as a count discrepancy if the respondent-reported household size was not blank and that was different from the number of data defined persons on the mail back short form. Long forms failed for this reason, but also failed if the respondent-reported household size **was** blank and both the number of data defined persons on the long form was less than six **and** the number of names on the roster was different from the number of data defined persons on the mail back long form. That extra comparison using the roster may have accounted for the greater failure rate of count discrepancy cases for long forms compared to short forms.

Another difference to note is that long form cases were almost a third less likely than short forms cases to be selected for large household followup. Since long forms were randomly assigned to households, we see that respondents in large households that received a long form were less likely to return their mailback form than large household that received the short form.

Respondents could only respond by Internet if they were in the short form universe and had their census ID number available. Therefore, as one might expect, Internet forms failed at rates very similar to short forms.

The rate for Be Counted forms needs some explanation. Only BCFs that were reporting a whole household were eligible for large household followup and no BCFs were eligible for count discrepancy followup. This count of BCFs, 598,994, include both whole household and partial household cases. Therefore, the rates for BCFs should not be compared directly to the other form types.

**Table 4. Coverage edit followup edit failure rate by form type and edit failure type**

Type of Form	Eligible Forms	All Coverage Edit Cases		CDFU Only		LHHFU Only	
		Number	Failure Rate	Number	Failure Rate	Number	Failure Rate
<b>Totals</b>	<b>82,008,049</b>	<b>2,544,072</b>	<b>3.1%</b>	<b>1,148,449</b>	<b>1.4%</b>	<b>1,395,623</b>	<b>1.7%</b>
Short Forms	69,235,695	2,022,196	2.9%	790,470	1.1%	1,231,726	1.8%
Long Forms	12,106,988	514,098	4.2%	357,369	3.0%	156,729	1.3%
Internet Forms	66,372	1,837	2.8%	610	0.9%	1,227	1.8%
Be Counted Forms	598,994	5,941	1.0%	n/a	n/a	5,941	1.0%

Source: CEFU input files

#### 4.1.5. Failure rates by edit failure reason

The number of edit failures by failure reason are shown in Table 5. Slightly more than half of all coverage edit followup cases were large household or possible large household cases. For the count discrepancy cases, more cases failed the coverage edit when there were more data defined persons than the respondent-reported household size.

**Table 5. Coverage edit followup failure rate by edit failure reason**

Type of Edit Failure	Coverage Edit Cases	
	Total Selected	Percent of Workload
<b>Total of all Addresses</b>	<b>2,544,072</b>	<b>100.0%</b>
Total Large Household Cases	1,395,623	54.9%
Definite Large Household	1,334,300	52.4%
Possible Large Household	61,323	2.4%
Total Count Discrepancies	1,148,449	45.1%
Count Discrepancy - High	699,379	27.5%
Count Discrepancy - Low	449,070	17.7%

Source: CEFU input files

#### 4.1.6. Failure rates vary by language of form

As shown in Table 6, there were 82,008,049 Census 2000 forms processed by June 8, 2000, that were eligible to be selected as CEFU cases. Of these, 2,544,072 forms or 3.1 percent were chosen as CEFU cases.

English language forms were much less likely to fail for coverage edit than non-English forms.

Non-English forms, including those in Spanish, Chinese, Vietnamese, Tagalog and Korean, failed coverage edit 10.7 percent of the time, compared to the rate for English language forms of 2.9 percent.

In Table 7, you can see that the non-English forms failed at greater rates for both CDFU and LHHFU cases. The non-English language cases failed for LHHFU three to five times as often as for English language forms while they failed CDFU at less than twice the rate as for English language forms.

**Table 6. Coverage edit followup edit failure rate by language of mailback form**

Language of Mailback Form	Eligible Forms	Coverage Edits	
	Number	Number	Failure Rate
<b>All Forms</b>	<b>82,008,049</b>	<b>2,544,072</b>	<b>3.1%</b>
<b>All English Forms</b>	<b>80,249,109</b>	<b>2,355,138</b>	<b>2.9%</b>
English (US)	80,245,150	2,349,029	2.9%
English (PR)	3,959	57	1.4%
other English*	n/a	6,052	n/a
<b>All Non-English Forms</b>	<b>1,758,940</b>	<b>188,934</b>	<b>10.7%</b>
<b>Total Spanish Language Forms</b>	<b>1,617,219</b>	<b>177,977</b>	<b>11.0%</b>
Spanish (US)	841,065	141,703	16.8%
Spanish (PR)	776,154	36,274	4.7%
<b>Total Asian Language Forms</b>	<b>141,721</b>	<b>10,957</b>	<b>7.7%</b>
Chinese	59,832	4,232	7.1%
Korean	39,254	1,882	4.8%
Tagalog	5,048	556	11.0%
Vietnamese	37,587	4,046	10.8%
Undetermined Asian Language Forms	n/a	241	n/a

\* includes all BCFs as well as 111 other forms where the mail back form type is unknown

Source: CEFU input files and DRF2

**Table 7. Coverage edit followup edit failure rate by language of mailback form and edit failure type**

Language of Mailback form	Coverage Edits		CDFU		LHHFU	
	Number	Failure Rate	Number	Failure Rate	Number	Failure Rate
<b>All Forms</b>	<b>2,544,072</b>	<b>3.1</b>	<b>1,148,449</b>	<b>1.4</b>	<b>1,395,623</b>	<b>1.7</b>
English Forms	2,355,138	2.9	1,104,298	1.4	1,250,840	1.6
Spanish Language Forms	177,977	11.0	41,054	2.5	136,923	8.5
Asian Language Forms	10,957	7.7	3,097	2.2	7,860	5.5

Source: CEFU input files

#### 4.1.7. Unsent cases not representative

Only a small number of cases, 85,183, were appropriately selected for the coverage edit universe but never sent to EDS for followup. They were checked in between May 15, 2000 and June 8, 2000. These cases made up only 3.3 percent of all coverage edit failure cases. However, the distribution of these cases among form types and language of mail back form varied widely from the distribution of the overall CEFU universe.

Table 8 shows a comparison of the CEFU failure rate by form type of cases that were sent to cases that were not sent. Cases selected from English language forms had a 3.2 percent chance of being in the unsent universe. Cases selected from Spanish language forms were less likely than cases selected from English language forms to be in the unsent universe. Only 1.6 percent of the forms selected from Spanish language forms were in the unsent universe.

More troubling is the fact that more than half of all the cases (53.8 percent) selected from Asian language forms were not sent for followup. Each of the cases from Asian language forms was translated or transcribed before being eligible to be selected for coverage edit followup. Apparently, this delayed the coverage edit selection for the majority of eligible cases from Asian language forms.

**Table 8. Likelihood coverage edit followup cases were not sent to the contractor by language of mailback form**

Type of Form	Total Eligible Forms Selected	Eligible Cases Not Sent	
		Number	Unsent Rate
<b>Totals</b>	<b>2,544,072</b>	<b>85,183</b>	<b>3.3%</b>
English Forms	2,355,138	76,382	3.2%
Spanish Forms	177,977	2,904	1.6%
Asian Language Forms	10,957	5,897	53.8%
Chinese	4,232	2,609	61.6%
Korean	1,882	768	40.8%
Tagalog	556	281	50.5%
Vietnamese	4,046	2,108	52.1%
Undetermined Asian Language Cases	241	131	54.4%

Source: CEFU input files and DRF2

Table 9 shows a comparison of the CEFU failure rate by form type of cases that were sent to cases that were not sent. While only 19.2 percent of the cases that were sent were long form cases, 48.0 percent of the forms not sent were long form cases. In the future, we need to be much more aware of the coverage implications of eliminating some of the coverage edit universe.

**Table 9. Coverage edit followup edit failure rate for cases sent to the contractor compared to cases not sent by form type of mailback form**

Type of Form	Total Forms Selected	Coverage Edit Cases Sent		Coverage Edit Cases Not Sent	
		Number	Percent of Sent Cases	Number	Percent of Unsent Cases
<b>Totals</b>	<b>2,544,072</b>	<b>2,458,889</b>	<b>100.0</b>	<b>85,183</b>	<b>100.0</b>
Short Forms	2,022,196	1,977,900	80.4	44,296	52.0
Long Forms	514,098	473,211	19.2	40,887	48.0
Internet	1,837	1,837	100.0	0	0.0
Be Counted	5,941	5,941	100.0	0	0.0

Source: CEFU input files

## **4.2. How successful were we contacting and completing cases?**

In this section, we will present analysis on the likelihood that we were able to contact and then complete a followup interview with cases within the CEFU universe. This section looks at the success rates of the telephone appending service, the number of attempts made to contact and to complete a case, how often callbacks were needed, as well as the distribution of final case dispositions. As appropriate, counts and rates are shown by edit failure type, edit failure reason, and type of form.

As detailed in the methodology section of this report, 2,338,420 cases are considered as being appropriately sent to the contractor and returned to us with a complete, incomplete, or interim case dispositions. The following analysis uses this as the universe of CEFU cases. The contractor, EDS, made attempts to contact and conduct a CEFU interview for all of these cases.

During interviewing hours, if a telephone interviewer was available to conduct an interview, the autodialer dialed new or incomplete cases at a certain ratio of calls to available interviewers. If there was no answer or an answering machine answered, the case was returned to the queue to be called again according to the calling strategy. If the call was answered, the call was routed to the next available interviewer. This dialing strategy minimized the number of times a respondent answered the phone but no interviewer was available to begin the interview. Unfortunately, this strategy also sometimes led to interviewers sitting around waiting for calls.

The telephone interviewer would begin by asking questions to determine if we had reached the correct household. If so, the interview would begin. If it was the correct household, we determined if there was an eligible respondent available. The person who completed Census 2000 mailback form was always eligible to respond to the CEFU interview. The second person listed on the form was also eligible if they were at least 18 years old. No one else was permitted to respond to the CEFU interview.

Our decision to only allow person one or person two to be eligible respondents was made to make it more likely that high quality data would be collected from a knowledgeable respondent for each CEFU interview. This was especially important when the collection of long form data was required during the CEFU interview. We also needed to ensure we avoided disclosure of Census data except to the respondents themselves. Unfortunately, this eligibility standard may have made it more difficult to complete cases.

There were other requirements placed on the contractor after the initial awarding of the contract that had a direct impact on their ability to contact cases and complete interviews. In addition to our eligibility standards and the conservative dialer settings, we also required that fifty percent of the contacts attempts be made on weekends to help us reach more respondents. This led to a need for greater staffing capacity than initially had been envisioned.

If these requirements had been defined up front at the time of contract award, the contractor may have been able to better accommodate the requirements and may have been more successful completing cases.

#### 4.2.1. Look up and telephone number quality

Every case selected for CEFU was initially sent to a telephone appending service for a first pass review. This service attempted to provide a telephone number, if the telephone number was missing, and attempted to correct any wrong numbers. Later, if the call center determined through contact with a household that the case had an incorrect telephone number, they returned the case for further research by the appending service.

It was found during production that this second pass by the appending service had a very limited success at identifying valid telephone numbers for these cases (3.9 percent success on a second try). The decision was made to not send additional bad telephone numbers to the appender a second time as of Friday, July 21, 2000. Therefore, 102,353 cases with a final status as non-interviews were not sent to the appender a second time. We assume that these cases were similar to the cases that were sent to the appender a second time and that their removal from percentages in Table 10 does not skew the distribution.

Table 10 shows the results of the telephone appending operation by final disposition. Of the cases that were sent to the appender, the appender provided a new or corrected telephone number 7.4 percent of the time (172,633 out of 2,338,420). Of the cases where telephone numbers were changed by the appender on either the first or second pass, 21.1 percent ended up being complete cases. The remainder, which were initially noninterviews and only sent to the appender one initial time, were all noninterviews.

The majority of the time a telephone number was changed by the appending service, we were still unable to complete that case. While 12.5 percent of all non-interview cases had a telephone number provided by the appending service, only 2.9 percent of all completed cases had the unit's telephone number changed by the appender. We should investigate whether there are better methods or services at obtaining missing or incorrect telephone numbers for our followup operations.

**Table 10. Results of the telephone appending operation by final disposition**

Telephone Number	Final Disposition Type				
	All Cases	Completed Cases		All Non-Interview Cases	
	Number	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>2,338,420</b>	<b>1,251,971</b>	<b>53.5%</b>	<b>1,086,449</b>	<b>46.5%</b>
Not Changed	2,063,434	1,215,571	58.9%	847,863	41.1%
Changed at Least Once	172,633	36,400	21.1%	136,233	78.9%
Not Sent to the Appender	102,353	0	0.0%	102,353	100.0%

Source: CEFU evaluation files

Table 11 shows the results of the telephone appending operation from cases with a final disposition of 'no valid telephone number'. There were 562,049 cases that had a final disposition of 'no valid telephone number'. The appending service was unable to offer an



alternative telephone number for 67.9 percent of these cases. For 78,288 cases, another telephone number was supplied on the first or second attempt (or both) which later turned out to also be an invalid telephone number(s). For the remainder, 18.2 percent of the cases were not sent to the appending service a second time.

**Table 11. Results of the Telephone Appending Operation for Cases with a Final Disposition of “No Valid Telephone Number”**

Telephone Number	All Cases	
	Number	Percent
<b>Total Number of Cases</b>	<b>562,049</b>	<b>100.0%</b>
Not Changed	381,408	67.9%
Changed at Least Once	78,288	13.9%
Not Sent the Appender	102,353	18.2%

Source: CEFU evaluation files

#### 4.2.2. Making contact

Calls for CEFU were made seven days a week from 8 am to 10 pm local time. If a case was not completed during the initial call, the case was called again until the maximum number of call attempts was reached. A maximum of twelve calls were made to each case to establish contact and determine if we had reached the correct household. These calls were spread across different days of the week and times of day. We also required that half (six of twelve) calls be made on weekends.

Valid data on the number of calls to complete cases or establish contact with the respondent were only available for cases completed through June 30, 2000. Cases returned after that date may have been reallocated to other call centers, causing them to have invalid values. This limitation minimizes the conclusions we can draw from the data in Tables 12-13.

Table 12 shows the distribution of calls per case to establish contact with the household for completed edit failure cases. A majority of cases that were completed (64.4 percent) only required one call to establish contact with the household. In fact, over 96 percent of all completed cases had contact established in four or fewer attempts. This distribution of calls per case to establish contact with the household was consistent across edit failure reasons.

**Table 12. Distribution of calls per case to establish contact with the household for completed edit failure cases returned prior to July 1, 2000**

Calls Per Case to Establish Contact with the Household	All CEFU Cases	
	Number	Percent
<b>Number of Cases</b>	<b>826,806</b>	<b>100.0%</b>
Exactly 1	532,103	64.4%
2	167,590	20.3%
3	65,079	7.9%
4	29,535	3.6%
5	14,539	1.8%
6 or more	17,960	2.2%

Source: CEFU evaluation HH level files - variable alltrys

We also attempted to collect data on how many calls were made to each case to complete it once contact had been established. Table 13 shows the distribution of calls per case after establishing contact with the household to complete each edit failure cases returned prior to July 1, 2000.

Clearly, very few cases were completed if they were not completed in the first few call attempts. No additional calls were required 78.2 percent of the time and over 95 percent of these cases required two or fewer additional call attempts. This distribution of calls per case after establishing contact with the household was consistent across edit failure reasons.

If we had an opportunity to test CEFU during the dress rehearsal for Census 2000, we may have been able to revise the number of call attempts required. This would definitely have had an impact on costs for CEFU in Census 2000. Also, making fewer call attempts would expedite the availability of unresolved cases for a potential personal visit followup.

**Table 13. Distribution of calls per completed case after establishing contact with the household for cases returned prior to July 1, 2000**

Duplicates Per Case	All CEFU Cases	
	Number	Percent
<b>Number of Cases</b>	<b>826,806</b>	<b>100.0%</b>
No Additional Calls	646,888	78.2%
Exactly 1	102,230	12.4%
2	39,295	4.8%
3	18,135	2.2%
4	9,040	1.1%
5	4,921	0.6%
6 or more	6,297	0.8%

Source: CEFU evaluation HH level files - variable intrtrys

For several reasons, nothing can be said about the calling distributions for non-interview cases. As stated above, these evaluation data file variables (INTTRYs and ALLTRYs) were filled incorrectly after July 1, 2000. All non-interview cases were returned to us in two files at the conclusion of the operation without a record of if, and if so when, calling had ceased for each case. The Decennial Statistical Studies Division (DSSD) had not asked the contractor to record a case level date when call attempts ceased. The information about the number of call attempts for non-interview cases, both prior to and after establishing initial contact with the household, would have been sufficient for our evaluation needs, had it been available for the entire length of the operation.

When phase two began at the beginning of August, all evaluation variables were reset, losing the history of the case. This was agreed to at the time by DSSD because we were told it was the only way they could initiate phase two. There is no way of knowing whether a non-interview case with three recorded call attempts made actually had more during the first phase of the operation - before the evaluation variables were reset. All non-interview cases where calling had ceased during phase one because the maximum number of call attempts had been reached should have been attempted again in phase two.

It would be valuable when conducting future research for us to collect and analyze data on the number of call attempts for us to establish contact with households as well as the number of attempts for us to complete cases in a telephone followup operation.

#### *4.2.3. Call backs needed*

During the course of an interview, a respondent could request that the interview be completed at a later time. The telephone interviewer would try to set a scheduled callback by asking the respondent for a time and date to complete the interview. If no time or date was provided, an

unscheduled callback was set. In either case, an alternative telephone number could be recorded for the callback. Unfortunately, evaluation variables carrying this information were reset when phase two began. Therefore, the data concerning noninterview cases reflect ONLY the callback attempts made after the start of phase two.

Table 14 shows the likelihood that at least one callback was set for each completed case. Respondents requested at least one callback in 38.3 percent of all completed CEFU cases. Respondents for large household cases were a little more likely to request a callback than respondents for count discrepancy cases.

**Table 14. Likelihood of callbacks for completed cases by type of edit failure**

	Edit Failure Type					
	Completed Cases		Count Discrepancy		Large Household	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>487,265</b>	<b>100.0%</b>	<b>764,706</b>	<b>100.0%</b>
No Callbacks	772,951	61.7%	328,726	67.5%	444,225	58.1%
One or more Callbacks	479,020	38.3%	158,539	32.5%	320,481	41.9%

Source: CEFU evaluation HH level files - variable ccbck

Table 15 shows the likelihood that at least one callback was set for each completed case by count discrepancy type. Among count discrepancy cases, those with low data defined count had a higher rate of callbacks (40.1 percent).

**Table 15. Likelihood of callbacks for completed cases by count discrepancy type**

	Count Discrepancy Type					
	Count Discrepancy		High Data Defined Person Count		Low Data Defined Person Count	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>487,265</b>	<b>-</b>	<b>322,509</b>	<b>-</b>	<b>164,756</b>	<b>-</b>
No Callbacks	328,726	67.5%	230,098	71.3%	98,628	59.9%
One or more Callbacks	158,539	32.5%	92,411	28.7%	66,128	40.1%

Source: CEFU evaluation HH level files - variable callback

Table 16 shows the likelihood that at least one callback was set for each case that was not completed. About one quarter of the incomplete CEFU cases were set for a callback to be made. Respondents for large household cases were more likely to request at least one callback than respondents for count discrepancy cases.

**Table 16. Likelihood of callbacks for incomplete cases by type of edit failure**

	Edit Failure Type					
	Incomplete Cases		Count Discrepancy		Large Household	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>1,086,449</b>	<b>100.0%</b>	<b>518,377</b>	<b>100.0%</b>	<b>568,072</b>	<b>100.0%</b>
No Callbacks	823,911	75.8%	414,985	80.1%	408,926	72.0%
One or More Callbacks	262,538	24.2%	103,392	19.9%	159,146	28.0%

Source: CEFU evaluation HH level files - variable ccbck

Table 17 shows the likelihood that at least one callback was set for each completed case by count discrepancy type. Among count discrepancy cases, those with low data defined count had a higher rate of callbacks (22.5 percent).

**Table 17. Likelihood of callbacks for incomplete cases by count discrepancy type**

	Count Discrepancy Type					
	Count Discrepancy		High Data Defined Person Count		Low Data Defined Person Count	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>518,377</b>	<b>-</b>	<b>321,804</b>	<b>-</b>	<b>196,573</b>	<b>-</b>
No Callbacks	414,985	80.1%	262,552	81.6%	152,433	77.5%
One or More Callbacks	103,392	19.9%	59,252	18.4%	44,140	22.5%

Source: CEFU evaluation HH level files - variable ccbck

The calling protocol for the CEFU program was to make up to twelve callbacks to a household after we had established it was the correct household and the respondent requested a callback to complete the interview. Over 99 percent of cases involving a callback took six or fewer callbacks to complete the case.

According to our evaluation data, approximately a quarter of the incomplete cases requested a callback. Of these, very few received the full twelve call back attempts. In fact, less than three percent of these cases received more than six call back attempts.

It should also be noted that while we conducted callbacks with respondents to complete the followup, we did not provide the ability for the respondent to call us back at their convenience. Some telephone surveys provide respondents with this ability. In the future, we should consider providing the ability for the respondent to call us back to complete their followup interview in order to potentially increase the completion rate and minimize the number of callback attempts.

#### 4.2.4. Completion of coverage edit followup cases

When the operation ended, each case was assigned a final disposition code. Table 18 presents a summary of the number and percent of the final disposition of all coverage edit followup cases. All final disposition codes can be grouped into one of four categories: complete cases, refusals, cases with no valid telephone number, and other non-interviews.

**Table 18. Number and percent of the final disposition of coverage edit followup cases**

	CEFU Cases	
	Number	Percent
<b>All Eligible Delivered Cases</b>	<b>2,338,420</b>	<b>100.0%</b>
<b>Total Complete Cases</b>	<b>1,251,971</b>	<b>53.5%</b>
Fully complete interview	1,028,207	44.0%
Sufficient partial interview	216,875	9.3%
Count complete interview	6,889	0.3%
<b>Refusals</b>	<b>201,385</b>	<b>8.6%</b>
<b>No Valid Telephone Number</b>	<b>562,049</b>	<b>24.0%</b>
<b>Other Non-Interviews</b>	<b>323,015</b>	<b>13.8%</b>
No contact after 12 call attempts	59,459	2.5%
Contact made but incomplete after 12 callback attempts	9,834	0.4%
Ineligible respondent only	1,858	0.1%
Contact made, but interview incomplete after < 12 callbacks*	8,927	0.4%
No contact after <12 call attempts*	170,919	7.3%
Case was never attempted*	72,018	3.1%

\* Interim non-interview codes can be misleading. When we implemented phase two of the CEFU operation, any cases with an interim disposition codes were 'reset'. These cases would have looked like newly delivered cases that were never attempted. We lost all the history of these cases. Therefore, most, if not all of the cases which ended up with interim codes probably were attempted in the first phase of the CEFU operation some number of times.

Source: CEFU evaluation HH level files - variable codecase

Completed cases were ones in which we were able to get responses to all nine coverage probe questions to the respondent. Overall, 53.5 percent of all valid cases were completed. Within the complete cases, there were three categories. Most cases (44.0 percent) were fully complete. This means we got responses for every census data question for each person for whom we needed to

collect information. Partially completed cases had responses collected for some but not all of the data needed for persons for whom we needed to collect information. This could mean as few as one question was unanswered for one person or only two questions per person were answered. Sufficient partial cases made up 9.3 percent of the universe of cases. A case was considered count complete if for one or more of the persons for whom we wanted to collect information, we were unable to collect any information beyond name. This occurred only in 0.3 percent of the cases.

Looking at the rest of Table 18 we see that 46.5 percent of all selected cases were not completed. The main reason for cases not being completed was our inability to get valid telephone numbers. In 24.0 percent of all CEFU cases, we neither captured a valid telephone number from the mail back Census form nor were we able to obtain a valid telephone number from the telephone appending service subcontracted by EDS. Research needs to be done to find better ways of ensuring we have a valid telephone number available to use for followup operations. This may be by doing a better job obtaining missing or invalid telephone numbers for future telephone followup operations or by improving the likelihood we collect a valid telephone number from the respondent in the first place.

Refusals were 8.6 percent of all cases. In these cases, we did contact an eligible respondent and that person refused to respond to our followup. This could have occurred initially or after answering some of the followup questions. At minimum, a respondent must have answered all of the questions through the nine coverage probe questions for the interview to be considered complete. We did not plan a refusal conversion operation during this operation. Research should be done into reasons for followup refusals so we can have more completed cases during future telephone followup operations.

Nearly 14 percent of cases were other non-interviews. When phase one of the coverage edit followup operation was concluded on July 30, 2000, most of the still incomplete cases were reset to a disposition of 99, indicating that they had never been attempted. Because phase two was not planned for from the start, we were unable to retain dispositions for incomplete cases prior to phase two. Due to limited development time, compromises were necessary to allow us to increase our completion rate by implementing phase two.

These non-interview cases were included in phase two of the CEFU operation. This means these reset cases could have been called as many as twenty-three times (a callback scheduled on the eleventh contact attempt, followed by 12 call back attempts), or never called at all, prior to the start of phase two. They could have already been refused. Or, it may have been that a callback was needed in Spanish (language difficulty cases for languages other than Spanish, including one of the four Asian languages, were not reset). Therefore, little should be concluded based on the differences between the different non-interview dispositions.

It is clear that we were unable to complete the interview before the end of phase two for the non-interview cases. For most of these cases, we did not make the maximum number of call attempts during phase two (although more calls may have been made for these cases during phase

one). This included cases where there was no contact after less than twelve call attempts (7.3 percent), where contact was made, but the interview was incomplete after less than twelve call back attempts (0.4 percent), where the case was never attempted (3.1 percent), and when only ineligible respondent were reached (0.1 percent). In about three percent of the cases, we did make the maximum number of call attempts during phase two and were still unable to complete the interview. In this situation, either there was no contact after a full twelve call attempts (2.5 percent) or that contact was made but the interview was incomplete after a full twelve call attempts (0.4 percent).

#### *4.2.5. Completion rate verses response rate*

There are several ways to define a completion rate for the CEFU operation. One way is to compare the number of cases we were successful with to the total number of eligible cases we chose for the CEFU universe. This follows one of the response rate definitions issued by the American Association of Public Opinion Research (AAPOR). It is defined as the number of complete interviews, including partial interviews, divided by the number of interviews (complete and partial) plus non-interviews (all non-interviews including refusals) plus all cases of unknown eligibility (cases with no valid telephone number). This minimum response rate is contained back in Table 18 and indicates a response rate of 53.5 percent.

This rate is appropriate if one was considering how successful we were getting additional information from all the cases we wanted to follow up. This rate considers cases we were unable to reach by telephone, for any reason, as nonrespondents. As we think about possible CEFU operations in future Decennial Censuses, this rate allows us to focus on both the effectiveness of completing cases as well as the appropriateness of the methods we chose to try and complete them.

Another way is to look at this is by removing cases of unknown eligibility from the denominator. To do this, we need to either estimate what proportion of cases of unknown eligibility (cases with a disposition of 'no valid telephone') are actually eligible or assume they are all ineligible. Doing this allows us to include just the cases where a telephone contact was possible in our calculations. The AAPOR also defines this type of rate. Such a rate is focused more on the effectiveness of our CEFU operation in completing cases that were actually possible to contact by telephone. The maximum value would be calculated if we assumed that all cases of unknown eligibility are ineligible. This maximum response rate, calculated according to AAPOR standards, is 70.5 percent .

This rate is much higher than the first rate but it is ignoring the fact that by choosing to make this a telephone only operation with no field followup, we made it impossible to interview a substantial portion of the universe (24.0 percent). In the future, we should consider ways to make it possible to reach the types of households we were unable to reach during the CEFU operation during Census 2000. Possible improvements may include improving methods at obtaining valid telephone numbers and/or conducting a field followup operation for cases we cannot reach by telephone.



4.2.6. *Some differences by edit failure type - large household followup cases verses count discrepancy followup cases*

There were some differences in the final disposition distribution between count discrepancy cases and large household cases. Table 19 shows that large household cases (57.4 percent) were more likely to be completed than count discrepancy cases (48.5 percent). These count discrepancy cases were almost twice as likely to result in a refusal (11.2 percent) than the large household cases (6.7 percent).

**Table 19. Number and percent of coverage edit followup cases by disposition category by edit failure type**

	All cases		Count Discrepancy		Large Household	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>2,338,420</b>	<b>100.0%</b>	<b>1,005,642</b>	<b>100.0%</b>	<b>1,332,778</b>	<b>100.0%</b>
Total Complete Cases	1,251,971	53.5%	487,265	48.5%	764,706	57.4%
Refusals	201,385	8.6%	112,522	11.2%	88,863	6.7%
Total Non-Interviews	323,015	13.8%	127,074	12.6%	195,941	14.7%
No Valid Telephone Number	562,049	24.0%	278,781	27.7%	283,268	21.3%

Source: CEFU evaluation HH level files - variable codecase

Table 20 shows the disposition rates by type of Census form submitted. Short forms and long forms, which make up over 99 percent of all cases, had nearly the same distribution of outcomes for the total and by edit failure type. The CEFU cases originally submitted over the Internet were more likely to result in a non-interview case than cases submitted on a short or long form. Also, respondents in large households who submitted their Census 2000 form over the Internet were much more likely to have reported valid telephone numbers or at least likely to be found in the telephone appending operation.

**Table 20. Percent of coverage edit followup completed cases by disposition category and type of census form submitted**

	Form Type of Census 2000 Form			
	Short Forms	Long Forms	Internet	Be Counted
Count Discrepancy Cases				
Total Complete Cases	47.8%	50.1%	46.4%	n/a
Refusals	11.4%	10.8%	8.8%	n/a
Total Non-Interviews	12.6%	12.6%	21.6%	n/a
No Valid Telephone Number	28.2%	26.6%	23.1%	n/a
Large Household Cases				
Total Complete Cases	57.2%	59.4%	58.0%	45.4%
Refusals	6.8%	5.3%	6.2%	10.5%
Total Non-Interviews	14.9%	12.9%	23.8%	19.9%
No Valid Telephone Number	21.1%	22.4%	12.1%	24.2%

Source: CEFU evaluation HH level files - variable codecase

One interesting note: there were concerns that the length of the CEFU interview, as well as the repetition of the nine coverage questions, would lead to a high refusal rate and hurt the completion rate. By this reasoning, we might expect that long form cases, which on average took longer than short form cases, would have lower completion rates and higher refusal rates than for short forms. As shown in Table 20, long forms for both large household cases and count discrepancy cases had higher completion rates and lower refusal rates than short forms.

Upon further reflection, when we began the CEFU operation, we did not conduct followup interviews with long form cases. This was both by design and necessity. We believed that allowing the interviewers to begin by conducting only the shorter, simpler short form interviews would allow them to become proficient with the screener questions and short form questions before tasking them with the longer and more difficult long form.

Additionally, at the start of the operation, the long form instrument had not been completely tested. When it was ready, we only allowed the five call centers with the best performance conducting short form cases to conduct CEFU for long form cases. At that point, the telephone interviewers at these five call centers had experience with the CEFU followup interview, including the coverage probes and the short form questions to collect person data. This indicates that the experience of the telephone interviewers and of the call centers with this program was much more important to an increased completion and a decreased refusal rate than the length of the followup interview. Also, the strategy of starting with only the short form interviews seems to have been successful.

4.2.7. *Big differences by language of form*

Respondents for Census 2000 had several choices of the language of the form they submitted to the Census Bureau. While the vast majority used an English language form, there were also forms available in Spanish and several Asian languages. There were two versions of the form in English and in Spanish - one for households in the United States (U.S.) and the other for those in Puerto Rico. Additionally, versions of the form were available in four Asian languages: Chinese, Korean, Tagalog, and Vietnamese. Table 21 shows the number of CEFU cases by the language of the Census 2000 form the respondent submitted.

**Table 21. Number of coverage edit followup cases by disposition category and language of census form**

Number of Cases	Language of Forms					
	All Forms	English (US)	English (PR)	Spanish (US)	Spanish (PR)	Asian language Forms
<b>All CEFU Cases</b>	<b>2,338,420</b>	<b>2,164,509</b>	<b>52</b>	<b>134,791</b>	<b>34,526</b>	<b>4,542</b>
Count Discrepancy Cases	1,005,642	966,660	25	20,925	16,621	1,411
Large Household Cases	1,332,778	1,197,849	27	113,866	17,905	3,131

Source: CEFU evaluation HH level files

Table 22 shows the distribution of disposition categories by language of census form. Forms submitted in both English and Spanish from Puerto Rico were much less likely to be completed compared to forms from the U.S. This was due primarily to our inability to obtain valid telephone numbers for almost sixty percent of the cases from Puerto Rico. While Spanish forms for U.S. responses had a rate of invalid telephone numbers of 19.5 percent, Spanish forms for responses from Puerto Rico had a rate of 58.3 percent. Also, English forms for U.S. responses had a rate of invalid telephone numbers of 23.8 percent compared to 59.6 percent of Spanish forms for responses from Puerto Rico.

The inability to get telephone numbers in Puerto Rico is related to the quality of the address list in Puerto Rico. During the address list development process, the addresses in Puerto Rico were location description types of addresses. This may have caused problems with getting the telephone number. It is also possible that a lower percent of Puerto Rican households have telephones than households in the U.S..

Also, non-English forms, including all forms submitted in Spanish as well as those in any of the four Asian languages (Chinese, Korean, Tagalog, and Vietnamese) were much less likely to result in refusals than English forms.

**Table 22. Percent of coverage edit followup cases by disposition category and language of census form**

Percent of Coverage Edit Followup Cases	Percent of Forms by Language					
	All Forms	English (US)	English (PR)	Spanish (US)	Spanish (PR)	Asian Language Forms
<b>Total cases</b>						
<b>Total Complete Cases</b>	<b>53.5%</b>	<b>53.3%</b>	<b>23.1%</b>	<b>62.3%</b>	<b>31.7%</b>	<b>52.8%</b>
<b>Refusals</b>	<b>8.6%</b>	<b>9.0%</b>	<b>7.7%</b>	<b>3.9%</b>	<b>1.2%</b>	<b>1.4%</b>
<b>Total Non-Interviews</b>	<b>13.9%</b>	<b>13.8%</b>	<b>9.6%</b>	<b>14.3%</b>	<b>8.8%</b>	<b>17.2%</b>
<b>No Valid Telephone Number</b>	<b>24.0%</b>	<b>23.8%</b>	<b>59.6%</b>	<b>19.5%</b>	<b>58.3%</b>	<b>28.5%</b>
Count Discrepancy Cases						
Total Complete Cases	48.5%	48.7%	24.0%	54.2%	27.8%	52.4%
Refusals	11.2%	11.5%	4.0%	4.2%	1.4%	1.6%
Total Non-Interviews	12.6%	12.6%	8.0%	15.5%	9.4%	16.8%
No Valid Telephone Number	27.7%	27.2%	64.0%	26.2%	61.4%	29.3%
Large Household Followup Cases						
Total Complete Cases	57.4%	57.1%	22.2%	63.8%	35.2%	53.0%
Refusals	6.7%	7.0%	11.1%	3.8%	1.1%	1.3%
Total Non-Interviews	14.7%	14.9%	11.1%	14.1%	8.3%	17.4%
No Valid Telephone Number	21.3%	21.0%	55.6%	18.3%	55.4%	28.3%

Source: CEFU evaluation HH level files

### **4.3. Were we successful improving coverage?**

The CEFU was designed to improve the within household coverage of a select universe of cases by probing the respondent to ensure that the household roster was correct. In 81.4 percent of all completed cases, no changes were made to the roster that was provided by the respondent on their mail back form. However, by reviewing the roster with the respondent and asking our nine probes questions, we increased our confidence in the accuracy of these 1,019,194 forms where no changes were made to the roster.

The other 232,777 completed cases (18.6 percent) involved roster changes. This involved some combination of adding and/or removing names. Name adds were appended to the end of the roster as listed on the respondent's mailback form. Names were removed from the roster both to persons who were not residents of the household (deletes) as well as names which represent the same person as another name on the roster (duplicates).

All of the name adds, deletes, or duplicates were done either in response to a specific coverage probe or through the interrupt functionality of the instrument. Each question targets a specific group of people we know we have difficulty enumerating correctly. These questions can be found in Appendix A.

#### *4.3.1. What was the net coverage gain by person?*

The net coverage gain is determined by taking the number of persons added during CEFU and subtracting the number of persons removed from a household roster during CEFU.

Table 23 shows the number of roster name corrections due to adds, deletes, or duplicates. The CEFU operation actually resulted in a net loss of 105,199 persons compared to the originally completed Census self response forms. However, while the net improvement to the census from the CEFU operation was a decrease in the population, it did improve the accuracy of Census 2000. The CEFU ensured that 410,565 people who would have been counted in the wrong place or not at all were counted in the correct household.

**Table 23. Number of roster name corrections  
- names added, deleted, or removed as  
duplicates - by method of removal**

	Number of CEFU Persons
Adds	152,683
Deletes	207,182
Duplicates	50,700
<b>Number of Persons with Corrected Roster Status</b>	<b>410,565</b>
<b>Net Coverage Gain</b>	<b>(105,199)</b>

*4.3.2. Multiple adds, deletes and duplicates in one case*

Respondents were able to add, delete, or indicate the presence of a duplicate for one or more persons during the CEFU interview. Most often, no changes were made. However, when changes were made, sometimes multiple changes were made. Table 24 shows the distribution of completed cases by the number of person adds by edit failure case type. In 8.5 percent of completed cases there were one or more persons added to the household roster during the followup. Count discrepancy cases (12.3 percent) were more than twice as likely to have added a person to the household roster as large household cases (6.0 percent).

**Table 24. Distribution of completed cases by the number of person adds by completed edit failure case type**

Adds Per Case	Edit Failure Type					
	Completed Cases		Count Discrepancy		Large Household	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>487,265</b>	<b>100.0%</b>	<b>764,706</b>	<b>100.0%</b>
0 Adds	1,145,754	91.5%	427,266	87.7%	718,488	94.0%
One or More Adds	106,217	8.5%	59,999	12.3%	46,218	6.0%
1 Add	76,936	6.1%	45,170	9.3%	31,766	4.2%
2 Adds	16,564	1.3%	8,780	1.8%	7,784	1.0%
3 Adds	6,557	0.5%	3,612	0.7%	2,945	0.4%
4 Adds	3,143	0.3%	1,547	0.3%	1,596	0.2%
5+ Adds	3,017	0.2%	890	0.2%	2,127	0.3%

Source: CEFU evaluation HH level files - variable cadd

Tables C1 and C2 in Appendix C show the number of cases with at least one added person broken out by the two edit failure reasons (count discrepancy and large household) and their subcategories. The two large household categories, large household and possible large household, behave very similarly. However, there is a big difference when we look in Table C1 in Appendix C at the person add rates for the two count discrepancy categories - high data defined person count (HDDP) and low data defined person count (LDDP). LDDP count discrepancy cases were about six times as likely (27.5 percent to 4.5 percent) to have a person added during the CEFU than HDDP count discrepancy cases.

Table 25 shows that about ten percent of all cases had one or more persons deleted from the household roster during the followup. Again, we see that changes were much more likely for count discrepancy cases (16.0 percent) than for large household cases (6.5 percent). These deletes corrected the roster for 127,121 households.

**Table 25. Distribution of completed cases by the number of person deletes by completed edit failure case type**

Deletes Per Case	Edit Failure Type					
	Completed Cases		Count Discrepancy		Large Household	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>487,265</b>	<b>100.0%</b>	<b>764,706</b>	<b>100.0%</b>
No Deletes	1,124,850	89.8%	409,517	84.0%	715,333	93.5%
One or More Deletes	127,121	10.2%	77,748	16.0%	49,373	6.5%
1 Delete	91,773	7.3%	59,298	12.2%	32,475	4.2%
2 Deletes	17,417	1.4%	10,262	2.1%	7,155	0.9%
3 Deletes	6,017	0.5%	3,141	0.6%	2,876	0.4%
4 Deletes	4,366	0.4%	2,164	0.4%	2,202	0.3%
5+ Deletes	7,548	0.6%	2,883	0.6%	4,665	0.6%

Source: CEFU evaluation HH level files - variable cdup

Tables D1 and D2 in Appendix D show the number of cases with at least one deleted person broken out by the two edit failure reasons (count discrepancy and large household) and their subcategories. In Table D1, we see that completed possible large household cases were slightly more likely to have at least one person deleted as the completed large household cases. Table D2 shows the person delete rates for the two count discrepancy categories - high data defined person count (HDDP) and low data defined person count (LDDP). The HDDP count discrepancy cases were twice as likely (19.2 percent to 9.6 percent) to have a person deleted during the CEFU than LDDP count discrepancy cases.

Overall, just 2.2 percent of cases contained a name identified as a duplicate. Table 26 shows how often a case contained a name that was deleted from the roster because the respondent realized that it represented the same person that another name on the roster represented. These are really just a special kind of name delete. This was over five times as likely to occur in a count discrepancy case than in a large household case.



**Table 26. Distribution of completed cases by the number of person duplicates by completed edit failure case type**

Duplicates per case	Edit failure type					
	Completed cases		Count Discrepancy		Large Household	
	#	percent	#	percent	#	percent
<b>Total number of cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>487,265</b>	<b>100.0%</b>	<b>764,706</b>	<b>100.0%</b>
0	1,224,649	97.8%	466,308	95.7%	758,341	99.2%
1 or more names	27,322	2.2%	20,957	4.3%	6,365	0.8%
1	18,925	1.5%	15,475	3.2%	3,450	0.5%
2	2,937	0.2%	2,327	0.5%	610	0.1%
3	1,452	0.1%	1,089	0.2%	363	0.1%
4	1,430	0.1%	938	0.2%	492	0.1%
5+	2,578	0.2%	1,128	0.2%	1,450	0.2%

Source: CEFU evaluation HH level files - variable cdup

Tables E1 and E2 in Appendix E show the number of cases with at least one duplicated person broken out by the two edit failure reasons (count discrepancy and large household) and their subcategories. In Table E1, we see that completed possible large household cases were twice as likely to have at least one person removed as a duplicate as for the completed large household cases. Table E2 shows the person duplicate rates for the two count discrepancy categories - high data defined person count (HDDP) and low data defined person count (LDDP). The HDDP count discrepancy cases were about three times as likely (5.6 percent to 1.8 percent) to have a person removed as a duplicate than LDDP count discrepancy cases.

The instrument also allowed a name that was added to later be deleted. This could have been due to one of the four specific delete probe questions or through the interrupt option. Overall, 6,913 names were added and then deleted. This represents 3.3 percent of all deletes. The rest of the deleted names were originally data defined persons or names on the continuation roster from the originally submitted Census 2000 form.

Almost two percent of the completed cases had names both added and deleted. More than half of those had one add and one delete.

### *4.3.3. What happened to cases that were not completed?*

Almost half of all CEFU cases were not completed. This may have occurred because we chose not to attempt the case (see section 4.1.7), because we reached the limit on the number of attempts we made (see section 4.2.2), or because we were unable to complete an active case before the end of the CEFU operation.

CEFU cases that were not completed were processed like other cases that were not selected for CEFU. The final population count for these cases was determined in the census process called the Primary Selection Algorithm (PSA). This process looks at all census forms returned for a housing unit and selects person and housing unit data from these returns.

Prior to the PSA, The Decennial Response File (DRF) creation process linked related census forms into census returns and set an expected household return population count for each census return. This count represented the expected household size for each census return. This expected count is not the final household size. It was an intermediate count that reflects the maximum possible count for each census form. When PSA selects a census return, the expected return population count set in this pre-PSA DRF2 process usually is the final household size for the census return. For the purposes of this study, the expected household size is computed basis of a single return and is computed in a manner consistent with PSA computation of the return level expected household size.

According to these processes, large household cases (household size greater than six) that were not completed during CEFU kept the household size indicated by the sum of the number of data defined persons captured from the form and the number of names on the continuation roster. This was done in order to not exaggerate household size with a respondent reported household size that was exceptionally large compared to the number of persons listed on the return. The impact of this methodology as applied to CEFU cases that were not completed appears to be minimal.

Count discrepancy cases that were not completed during CEFU also went through this process, but the final household size was not determined the same way. In almost all cases, the maximum of the respondent reported household size and the number of data defined persons was assigned as the expected household size. However, this method was not uniformly successful in mimicking the final household size as determined by CEFU for both kinds of count discrepancy cases.

Count discrepancy cases where there were fewer data defined persons than the reported household size (or low data defined persons (LDDP)) were one type. For LDDP cases CEFU did complete, there were added persons only about a quarter of the time. For LDDP cases CEFU did not complete, the algorithm assigned the final household count using the maximum count - in this case the respondent reported household size - essentially adding people to the number of data defined persons every time. Assuming the cases that were not completed were behaving

similarly to those that were completed, we assigned the ‘wrong’ count 75 percent of the time. This resulted in a larger household size.

Count discrepancy cases where there were more data defined persons than the reported household size (or high data defined persons (HDDP)) were the other type. For HDDP cases CEFU did complete, there were deleted persons only about a quarter of the time. For HDDP cases CEFU did not complete, the algorithm assigned the household count using the maximum count, in this case the number of data defined persons on the form - essentially not deleting anyone. Assuming the cases that were not completed were behaving similarly to those that were completed, we assigned the ‘wrong’ count 25 percent of the time. This resulted in a larger household size.

This process of determining a household size for cases that were not completed during CEFU did not make an attempt to mimic the results of CEFU cases that were able to be completed. In the future, we should consider ways to increase the completion rate for CEFU cases to minimize the impact of assigning a final household size without additional information. We should also consider a new algorithm to assign a final household size for count discrepancy cases chosen for CEFU but not completed. This new method should more closely mimic the results for completed count discrepancy cases than the current method does.

#### **4.4. What were other characteristics of the interview process?**

The coverage edit followup instrument was effective in its two main objectives: correcting incorrect rosters and collecting person data. However, some desired functionality was not available. Unfortunately, CEFU was not developed as a proposal for Census 2000 until after the dress rehearsal in an effort to address coverage concerns. Therefore, we did not have the opportunity to test and improve the operation by conducting it in a census-type environment prior to Census 2000.

Both CEFU types, count discrepancies and large household cases, used the same strategy and procedures when contacting a household. This section will analyze several aspects of the CEFU instrument. We will discuss the effectiveness of the coverage probes, roster changes, and person data collection.

##### *4.4.1. Add probes*

Each respondent was asked five coverage probes designed to ensure consideration of several types of persons who may have been left off their household roster. In addition, respondents had the option to interrupt the interview to make further additions to the roster. Table 27 shows the number and percent of names added to the roster by each coverage probe. A total of 152,683 persons were added to the household roster during the CEFU followup. More than half (54.7 percent) were added through the interrupt function. This accounted for 83,497 persons who were added to the roster of the appropriate household. We did not record any details about the reasons for these interruption adds during the interview.

The specific coverage probes targeting children and relatives were most successful, with 33,246 children and 17,088 relatives added during the CEFU interview. Probes targeting non relatives, persons staying temporarily or in the process of moving, and with a second residence were less successful, adding 10,442, 3,440, and 4,436 persons respectively. A small number of added persons, 534, were missing information on the evaluation file as to method of their addition.

**Table 27. Number and overall percent of names added to the roster by coverage probe**

	Added Persons	
	Number	Percent
<b>Total number of added persons</b>	<b>152,683</b>	<b>100.0%</b>
Children	33,246	21.8%
Relatives	17,088	11.2%
Non Relatives	10,442	6.8%
Temp Residents / Movers	3,440	2.3%
Second Residence	4,436	2.9%
Interruptions	83,497	54.7%
Undesignated Adds	534	0.4%

Source: CEFU evaluation person level files

### Following the Flow in Each Probe

Each of the add coverage probes had a similar three step flow - starting more broadly and then narrowing down the scope before confirming the addition of a name to the roster. Tables C3 - C13 (odd numbers) in Appendix C of this report show how often a complete case made it through each of the three steps of each coverage probe. There are separate tables for each of the five add probe questions as well as the add interrupt option. The data presented in these tables are limited to the 1,251,971 completed CEFU interviews. For a case to be considered complete, all of the nine coverage probe questions must have been asked and answered by the respondent. Also, there is an ordering effect present since the questions were always asked in the same order.

To begin the process of asking the coverage probes, the household roster as listed on the mailback form was read to the respondent. Then the coverage probes were asked one at a time. In the first step, the interviewer asked if a person with particular characteristics (child, roommate, and so forth.) was living or staying there around the beginning of April and was not included on that roster. The tables show how often this happened.

If the respondent said ‘no’, the interviewer moved on to the next coverage probe. If ‘yes’, the second step for the probe was to ask for that person’s name. If a name was offered by the

respondent, the interviewer reviewed the roster to determine if the same or a similar name was already listed. If necessary, the interviewer verified with the respondent that the similar names represent different persons. If the interviewer determined that the respondent believed the name was not already on the roster, the name was entered into the followup interview instrument. The tables show how likely a name was entered after a respondent answered 'yes' to the coverage probe.

If a name was entered, we moved on to the third step. In the third step, we confirmed with the respondent that this person **“was living or staying there most of the time as of April 1”**. This step explicitly informs the respondent about the residency rule. The tables show how likely a name was actually added to the roster after a name was entered in step two. This multi stage approach allowed the respondent to consider more possible residents while we defined the complete residency criteria within the final confirmation step.

In addition, the respondent could interrupt the interview at any point to add names to the household roster. Telephone interviewers would select the interrupt option and indicate that the respondent wants to add a person's name to the roster. The interviewer was directed to question the respondent about the reason for adding this person. The interviewer, based on their knowledge of the Census residence rules, was to determine if the person should indeed be added. Upon the completion of this action, the interview is resumed where it left off.

### **Comparing Across the Probes**

Tables C4 - C14 (even numbered only) in Appendix C of this report contain the same information on each of the coverage probes broken out by edit failure reason. What the variety of rates across coverage probe and edit failure reason shows is that these differences lead to very different likelihoods of a name actually being added to the roster. This may indicate that some of our residence rules are more intuitive to the respondents than others or that the coverage probes were not uniformly effective at presenting the residence rules.

Looking across the five add questions, those responding 'yes' to adding a person varied from 2.6 percent (questions 1 about children) to 0.7 percent (both question 4 about people temporarily away or moving and question 5 about people with no other permanent place to stay). Of those responding 'yes', the likelihood of a name being offered by the respondent varied from 87.7 percent of the time (question 2 about relatives) to 63.6 percent of the time (question 5 about people with no other permanent place to stay). Finally, of those cases where a name was given, the likelihood of a person actually being added varied from 87.0 percent (question 2 about relatives) to 51.0 percent (question 4 about people temporarily away or moving).

The interrupt option had the greatest impact on adding persons during the CEFU interviews. The flow of this option varied from the add coverage probes in that it was only a two step process. The interviewer selected the interrupt option to add a person in 5.2 percent of the cases. In 97.7 percent of the cases where the interrupt option was selected to add a person there was an actual person added.

## Added Persons by Edit Failure Reason

Table 28 shows the percent of complete cases with added persons broken down by probe question and edit failure reason. Here we see several interesting differences. The two types of large household cases were very similar in the likelihood of adding names from any one of the probes questions. However, the two types of count discrepancy cases were quite different.

Count discrepancy cases that resulted from the count of data defined persons exceeding the respondent-reported household size, known as high data defined person cases (HDDP), were about half as likely to add a person as the overall percentage for all cases. This relationship holds for all five add questions as well as the interrupt add option. This seems to make sense. These cases often had the correct number of data defined persons already on the form, with the respondent-reported household size being too low, or had one or more incorrect data defined persons that were deleted during CEFU, with the respondent-reported household size being correct.

Count discrepancy cases that resulted from the respondent-reported household size exceeding the count of data defined persons, known as low data defined person cases (LDDP), were two to four times as likely to add a person as the overall percentage for all cases for all probes except the one concerning persons temporarily away. This also seems to make sense. If the respondent-reported household size was actually the correct household size, then names had to be added to the roster during CEFU.

**Table 28. Percent of complete cases with adds by probe question and edit failure reason**

Question	All Complete Cases	Percent of Completed Cases			
		Count Discrepancy		Large Household	
		HDDP	LDDP	Definite	Possible
#1 Children	1.81%	0.82%	5.32%	1.46%	1.64%
#2 Relatives	1.06%	0.50%	2.22%	1.05%	1.07%
#3 Non relatives	0.68%	0.39%	1.66%	0.59%	0.72%
#4 Persons temporarily away	0.25%	0.12%	0.30%	0.29%	0.34%
#5 Persons with no other permanent Place to Stay	0.29%	0.17%	0.46%	0.31%	0.36%
Interruption option	5.05%	2.87%	19.15%	2.92%	3.10%
<b>Totals*</b>	<b>8.48%</b>	<b>4.54%</b>	<b>27.53%</b>	<b>6.02%</b>	<b>6.69%</b>

HDDP - high count of data defined persons compared to the reported household size

LDDP - low count of data defined persons compared to the reported household size

\* The totals include the 534 added persons for which the probe was not recorded. Columns do not sum because some cases had multiple adds due to different probes.

#### 4.4.2. Delete probes

Each of the delete coverage probes had a three step flow, similar to that of the add coverage probes. Table 29 shows the number and percent of names deleted from the roster by each coverage probe. A total of 207,182 persons were deleted from household rosters during the CEFU followup. More than three quarters were deleted through the interrupt function. This accounted for 164,368 persons who were removed from the roster of a household where they did not belong. We assume that this was usually done after the initial reading of the roster to the respondent.

The specific coverage probes targeting college students and persons with a second residence were most successful. The CEFU removed 19,103 college students and 16,255 persons with another residence where they lived or stayed most of the time. The two probes about members of the military and about persons in institutions such as prison, jail, mental hospitals or nursing homes, were less successful. Only 2,022 military members and 5,165 persons in institutions were removed from household rosters during CEFU.

**Table 29. Number and overall percent of names deleted from the roster by coverage probe**

	Deleted Persons	
	Number	Percent
<b>Total number of deleted persons</b>	<b>207,182</b>	<b>100.0%</b>
College Students	19,103	9.2%
Military	2,022	1.0%
Institution	5,165	2.5%
Second Residence	16,255	7.9%
Interruptions	164,368	79.3%
Undesignated Deletes	269	0.1%

#### Comparing Across the Probes

Tables D3 - D11 (odd numbers) in Appendix D of this report analyze the effectiveness of each of the four delete probe questions as well as the delete interrupt option at removing persons from the respondent-reported household roster. The data presented in these tables are limited to the 1,251,971 completed CEFU interviews. For a case to be considered complete, all of the nine coverage probe questions must have been asked and answered by the respondent. Also, there is an ordering effect present since the questions were always asked in the same order.

Tables D4 - D12 (even numbers) in Appendix D of this report contain the same information on

the first four delete probes and delete interrupt option, respectively, broken out by edit failure reason.

Looking across the four delete questions, those responding ‘yes’ varied from 2.8 percent (question 9 about having another residence) to 0.5 percent (question 7 about military members). Of those responding ‘yes’, the likelihood of a name being offered by the respondent varied from 81.3 percent of the time (question 9 about having another residence) to 73.2 percent of the time (question 7 about military members). Finally, of those cases where a name was given, the likelihood of a person actually being deleted varied from 81.8 percent (question 8 about institutions) to 40.7 percent (question 9 about having another residence).

The interrupt option had the greatest impact on deleting persons during the CEFU interviews. The flow of this option varied from the delete coverage probes in that it was only a two step process. The interviewer selected the interrupt option to delete a person in 8.2 percent of the cases. In 92.6 percent of the cases where the interrupt option was selected to delete a person there was an actual person deleted.

### **Deleted Persons by Edit Failure Reason**

Table 30 looks at the percentage of complete cases with deleted persons broken down by probe question and edit failure reason. Here we see several interesting differences. The two types of large household cases were mostly similar in the likelihood of deleting names from any one of the probes questions. However, the two types of count discrepancy cases were quite different.

Count discrepancy cases that resulted from the count of data defined persons exceeding the respondent-reported household size, known as high data defined person cases (HDDP), were more likely to delete a person than the overall percentage for all cases. This relationship holds for all four delete questions as well as the interrupt delete option. This seems to make sense. These cases usually had the correct number of data defined persons already on the form, with the respondent-reported household size being too low, or had one or more incorrect data defined persons that were deleted during CEFU, with the respondent-reported household size being correct.

Count discrepancy cases that resulted from the respondent-reported household size exceeding the count of data defined persons, known as low data defined person cases (LDDP), were about half as likely to delete a name from the household roster as the overall percentage for all cases. This also seems to make sense. These cases usually had too few data defined persons already on the form, with the respondent-reported household size being correct, leading to added persons, not deleting them.



**Table 30. Percent of complete cases with deletes by probe question and edit failure reason**

Question	All Complete Cases	Percent of Completed Cases			
		Count Discrepancy		Large Household	
		HDDP	LDDP	Definite	Possible
#6 College Students	1.32%	2.29%	0.64%	1.04%	1.44%
#7 Military	0.15%	0.21%	0.10%	0.14%	0.16%
#8 SP / GQ	0.38%	0.58%	0.20%	0.34%	0.34%
#9 Second Residence	0.94%	1.46%	0.54%	0.80%	0.96%
Interruption Option	7.55%	14.96%	8.21%	4.24%	5.26%
<b>Totals*</b>	<b>10.15%</b>	<b>19.23%</b>	<b>9.55%</b>	<b>6.39%</b>	<b>7.96%</b>

HDDP - high count of data defined persons compared to the reported household size

LDDP - low count of data defined persons compared to the reported household size

\* The totals include the 269 deleted persons for which the probe was not recorded. Columns do not sum because some cases had multiple deletes due to different probes .

*4.4.3. Names removed from the roster as duplicates*

The functionality to remove a name from the household roster that represented the same person as another name on the roster was handled two ways. At any time during the interview, but most likely during the initial reading of the roster, the respondent could have interrupted the interviewer and indicated which names were duplicates. The other opportunity was at the end of the followup interview. The interviewer looked over the roster (without reading it to the respondent) and thought about whether there might be duplicate names. If so, they asked the respondent if there were duplicates. If so, any duplicate names were removed from the roster.

In either case, there was only a one-step question flow. Once the respondent indicated that two or more names represented the same person, the telephone interviewer marked the name(s) to be removed. The interviewer was instructed to delete the name(s) for whom we were less likely to have already collected demographic data. The roster was then updated to reflect the removal of the duplicate(s).

Table 31 shows the number and percent of names deleted from the roster which represent the same person as another roster name by method of removal. A total of 50,700 persons were removed from the household roster because they represented the same person listed elsewhere on the roster. More than three quarters were removed as duplicates through the interrupt function. The remainder were removed during the final roster review.

**Table 31. Number and percent of names deleted from the roster which represent the same person as another roster name by method of removal**

	Duplicated Persons	
	Number	Percent
<b>Total number of duplicated persons</b>	<b>50,700</b>	<b>100.0%</b>
During Final Roster Review	11,235	22.2%
During Interruption	39,419	77.8%
Undesignated Duplicates	46	0.1%

**Comparing the two duplicate methods**

Tables E3 and E5 in Appendix E of this report analyze the effectiveness of two ways of removing persons from the respondent-reported household roster that represent the same person as another name on the roster, broken out by edit failure reason. The data presented in these tables are limited to the 1,251,971 completed CEFU interviews. Tables E4 and E6 in Appendix E of this report contain the same information on the two ways of indicating a duplicate, broken out by edit failure reason. Overall, count discrepancy cases were more than twice as likely to indicate the presence of a duplicate on the household roster for both duplicate methods as cases for any other edit failure reason.

Table 32 analyzes the effectiveness of two ways of removing persons from the respondent-reported household roster that represent the same person as another name on the roster, broken out by probe method and edit failure reason. Count discrepancy cases that resulted from the count of data defined persons exceeding the respondent-reported household size, known as high data defined person cases (HDDP), were far more likely to indicate the presence of a duplicate on the household roster as cases for any other edit failure reason.

**Table 32. Percent of complete cases with duplicates by probe question and edit failure reason**

Question	All Complete Cases	Percent of Completed Cases			
		Count Discrepancy		Large Household	
		HDDP	LDDP	Definite	Possible
After Last Coverage Probe	1.79%	4.78%	1.41%	0.59%	1.32%
Interruption Option	0.39%	0.81%	0.40%	0.21%	0.42%
<b>Totals</b>	<b>2.18%</b>	<b>5.58%</b>	<b>1.81%</b>	<b>0.79%</b>	<b>1.73%</b>

HDDP - high count of data defined persons compared to the reported household size

LDDP - low count of data defined persons compared to the reported household size

#### 4.4.4. Name edits

The CEFU instrument allowed the correction of names on the roster to help the telephone interviewer conduct the followup interview. As prescribed by DSCMO, these corrected names were not returned to the Census Bureau and were not reflected in Census 2000. However, as you can see in Table 33, the telephone interviewer needed to change one or more names in almost 30 percent of all completed CEFU cases. Data capture errors may have been the cause of many of these incorrect names.

**Table 33. Distribution of cases with name edits for completed edit failure cases**

Name Edits Per Case	All CEFU Cases		Edit Failure Type			
			Count Discrepancy		Large Household	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>487,265</b>	<b>100.0%</b>	<b>764,706</b>	<b>100.0%</b>
No Name Edits	878,150	70.1%	362,030	74.3%	516,120	67.5%
One or More Name Edits	373,821	29.9%	125,235	25.7%	248,586	32.5%
Exactly 1 Name Edited	186,782	14.9%	72,751	14.9%	114,031	14.9%
2 Names	91,105	7.3%	33,175	6.8%	57,930	7.6%
3 Names	30,814	2.5%	7,987	1.6%	22,827	3.0%
4 Names	22,846	1.8%	6,490	1.3%	16,356	2.1%
>5 Names	42,274	3.4%	4,832	1.0%	37,442	4.9%

Source: CEFU evaluation HH level files - variable *edit*

Table 34 shows that large household cases were a little more likely to have needed names edited than count discrepancy cases.

**Table 34. Percent of complete cases with at least one name edit by edit failure reason**

	Percent of Completed Cases				
	All Complete Cases	Count Discrepancy		Large Household	
		HDDP	LDDP	Definite	Possible
One or More Names Edits	29.9%	24.8%	27.4%	32.6%	31.6%

HDDP - high count of data defined persons compared to the reported household size

LDDP - low count of data defined persons compared to the reported household size

#### *4.4.5. Characteristics of the persons enumerated, deleted, or removed as duplicates in coverage edit followup*

The demographic profiles of persons on forms selected for CEFU varied in many ways from the overall population. The following section will look at the demographics for:

- Persons added during the CEFU interview
- Persons deleted during the CEFU interview or removed as duplicates.

There were 152,683 persons added to household rosters and 257,882 persons removed from (deleted or removed as duplicates) through CEFU. For these persons, we will look at frequencies and percent by sex, age, race, and Hispanic Origin. Also, household tenure will be compared between households chosen for CEFU and the overall household universe. Tables G1-G5, in Appendix G of this report, compare the distributions for each of these attributes for persons added to or removed from (deleted or marked as duplicates) household rosters during the CEFU interview to persons in the overall Census 2000 population. Note that the data for persons in the overall Census population were based on edited data, while the data in this report for persons enumerated or removed during CEFU were based on unedited data.

#### **Tenure**

There is very little difference between the distribution of owners and renters from CEFU cases with added or removed persons compared to the overall Census population.

#### **Sex**

Persons added through CEFU were slightly more likely to be male than persons in the overall Census population. Persons removed from household rosters through CEFU were slightly more likely to be female than persons in the overall Census population. Therefore, in regard to sex, the added and removed respondents were very similar to the overall Census population.

#### **Age**

Persons added through CEFU were much more likely to be 0-24 years in age than persons in the overall Census population. This is important because younger persons have been traditionally undercounted in the Census.

There were 257,882 persons removed through CEFU by deleting or being marked as duplicates. We had age data for less than half (96,209) of these persons. Those persons we had age data for were much more likely to be in the 15-24 or 85 and over age categories than persons in the overall Census population.

## **Race**

Persons added through CEFU were much more likely to have responded that they are Black, Native Hawaiian or other Pacific Islander, or some other race than persons in the overall Census population. This is important because minorities have been traditionally undercounted in the Census.

Persons removed from household rosters through CEFU were more likely to be Black than persons in the overall Census population.

## **Hispanic Origin**

Persons added through CEFU were much more likely to have responded that they are Hispanic than persons in the overall Census population. These added persons were more than three times as likely to be Mexican than the overall Census population. Persons responding as Other Spanish/Hispanic were almost twice as likely as the overall Census population. This is important because Hispanics have been traditionally undercounted in the Census.

Persons removed from household rosters through CEFU were mostly similar to persons in the overall Census population. The only exception is that removed persons were twice as likely as the overall Census population to be Cuban.

Additionally, there was a strong relationship between those who answered ‘Some other race’ to the race question and the results of the Hispanic Origin question. Almost 90 percent of those responding ‘other race’ also chose one of the Hispanic origin categories (the remaining 10 percent responded ‘not Hispanic’ or did not answer the question). Of those responding Mexican or Other Spanish/Hispanic to the Hispanic origin question, more than half responded ‘other race’ to the race question. Of those responding Puerto Rican to the Hispanic origin question, over 40 percent responded ‘other race’ to the race question.

It appears that Hispanic respondents may be much more likely to answer ‘other race’ to the race question when asked over the telephone during the CEFU interview compared to Hispanics responding to race question on the paper mailback form for Census 2000. Because the race and Hispanic origin data used in this evaluation are unedited, direct comparisons with published Census 2000 results are not possible. Further research should be done to see if this relationship, and possible mode effect, exists on the paper form for Census 2000.

### *4.4.6. What were the characteristics of the persons on the continuation rosters of large household cases?*

In addition to resolving count discrepancy cases, the other objective of the CEFU operation was to collect person data for persons on the continuation roster of large household cases. There were 1,327,756 persons listed on the continuation roster of large household cases which we collected

demographic data for during CEFU. In addition, there were 60,729 persons listed on continuation rosters for count discrepancy cases which were not large households.

The following section will only look at the demographics for persons enumerated during CEFU that were listed on the continuation roster of large household cases compared to persons in those same households that were data defined on the mail back Census form.

For these persons, we will compare frequencies and percent by sex, age, race, and Hispanic Origin. Also, in Table H1, in the Appendix of this report, household tenure will be compared between households chosen for CEFU and the overall household universe. Tables H2-H5, in Appendix H of this report, compare the distributions for each of these attributes for persons on the continuation rosters of large household cases to two other universes. Both the persons on Census forms where data were provided as well as the overall Census 2000 population are provided for comparison. Note that the data for persons in the overall Census population were based on edited data, while the data for persons enumerated during CEFU were based on unedited data.

For example, let's think about a typical Census 2000 form representing a household with eight persons. The respondent would have provided demographic data for the first six persons listed on the form. The last two would only have their names listed on the continuation roster. In Tables H2-H5, the first six persons in this household would be included in the column title "data defined persons from completed LHH cases (#1-6)". The last two persons listed would be included in the column titled "Persons from LHH continuation rosters (#7-12)."

### **Tenure**

Completed large household cases were more likely to be owners than the incomplete large household cases as well as the overall Census household population.

### **Sex**

Persons listed on the continuation roster and enumerated during CEFU/LHHFU were slightly more likely to be male than persons in the overall Census population. Persons that were data defined on forms that were completed during CEFU/LHHFU were only slightly more likely to be male than persons in the overall Census population.

### **Age**

Persons listed on the continuation roster and enumerated during CEFU/LHHFU were much more likely to be in the 0-14 age category than persons in the overall Census population. Persons that were data defined on forms that were completed during CEFU/LHHFU were much more likely to be in the 5-19 age category than persons in the overall Census population.

## **Race**

Persons listed on the continuation roster and enumerated during CEFU/LHHFU were much more likely to be reported as Some Other Race or Native Hawaiian or Other Pacific Islander than persons in the overall Census population. They were also much less likely to be reported as White.

Persons that were data defined on forms that were completed during CEFU/LHHFU were much more likely to be reported as Black than persons in the overall Census population. They were also much less likely to be reported as Some Other Race.

Of note: for completed LHHFU cases, data defined persons (#1-6) left the race question blank 27 percent of the time, but the continuation roster persons associated with those forms only left race blank 8 percent of the time. Those seem to be reporting 'Some other race' much more than expected. The persons in these households are mostly Hispanic, specifically Mexican.

## **Hispanic Origin**

Persons listed on the continuation roster and enumerated during CEFU/LHHFU were much more likely to have responded that they are Hispanic than persons in the overall Census population. This hold for all Hispanic Origin groups except Cuban and Puerto Rican.

Persons that were data defined on forms that were completed during CEFU/LHHFU were much more likely to have responded that they are Hispanic than persons in the overall Census population. This also holds for all Hispanic Origin groups except Cuban and Puerto Rican.

Additionally, there was a strong relationship between those who answered 'other race' to the race question and the results of the Hispanic Origin question. Over 90 percent of those persons listed on the continuation roster and enumerated during CEFU/LHHFU responding 'other race' (24.8 percent of continuation roster persons) also chose one of the Hispanic origin categories (the remaining 15 percent responded 'not Hispanic' or did not answer the question). Of those responding Mexican or Other Spanish/Hispanic to the Hispanic origin question, over 70 percent responded 'other race' to the race question. Of those responding Puerto Rican to the Hispanic origin question, over 40 percent responded 'other race' to the race question.

Over 85 percent of those persons that were data defined on forms that were completed during CEFU/LHHFU responding 'other race' (0.6 percent of data defined persons) also chose one of the Hispanic origin categories (the remaining 15 percent responded 'not Hispanic' or did not answer the question). However, of those responding anything to the Hispanic origin question, less than two percent responded 'other race' to the race question. More than half of those responding Mexican or Other Spanish/Hispanic, and over 405 of those responding Puerto Rican, left the race question blank when filling out their Census form.

It appears that Hispanic respondents may be much more likely to answer 'other race' to the race question when asked over the telephone during the CEFU interview compared to Hispanics

responding to race question on the paper mailback form for Census 2000. Because the race and Hispanic origin data used in this evaluation are unedited, direct comparisons with published Census 2000 results are not possible. Further research should be done to see if this relationship, and possible mode effect, exists on the paper form for Census 2000.

#### *4.4.7. Characteristics of households without a valid telephone number*

Tables I1 - I3 in Appendix I show details about the household and householder (person 1 was usually the person who filled out the form) for CEFU households with a final disposition of 'no valid telephone number'. It includes tenure of the housing unit and person demographic characteristics race and Hispanic origin, respectively, for the householder for cases where 'no valid telephone number' was the final disposition. These tables contain comparisons to the overall CEFU universe and the Census 2000 population. Note that the data for persons in the overall Census population were based on edited data, while the data for householders without a valid telephone number in CEFU were based on unedited data.

These CEFU households with a final disposition of 'no valid telephone number' were more likely to be households where the householder was a member of traditionally undercounted groups than those where we were able to get a valid telephone number. These cases were more likely to be renters, a traditionally undercounted group. The householder in these cases was more likely to be Black or American Indian, more likely to be of Hispanic Origin, and less likely to be white or Asian than persons within the CEFU universe of cases or the overall Census population.

Since there was no field followup component of CEFU, cases with a final disposition of 'no valid telephone number' could be not completed. By all three of these measures (tenure, householder's race and householder's Hispanic origin), the type of cases we were unable to contact were more likely to be members of traditionally undercounted groups. If we were able to follow up on these cases, we could potentially have done more to address the differential undercount in Census 2000. Therefore, for the next Census, we should consider conducting a field followup operation for cases where we are unable to obtain a valid telephone number.

#### *4.4.8. Length of calls*

The evaluation data for length of call were not correctly recorded. The contractor provided some information about call length for cases completed through July 31, 2000 - the end of the initial phase of CEFU. A sample of complete cases was taken by the contractor. Only cases that did not involve a callback were considered for this sample. Sample means, medians, and standard deviations were calculated for this sample of selected cases by the contractor. Table 35 shows the best information we have concerning length of calls.

It appears that call length means were close to our stated estimates about the average length of a coverage edit followup call. However, these call lengths do not include cases where more than



one attempt was necessary to complete the case. About two thirds of all complete cases required more than one contact attempt and these time estimate do not represent any of those cases.

**Table 35. Central measures of interview length by edit failure type and form type**

Type of Form and Edit Failure Type	Complete CEFU Cases Interviewed on the First Attempt			
	Estimated Interview Length (Minutes)	Mean Interview Length	Median Interview Length	Total Sample Selected
Short Forms (SF, IDC, BCF)				
Count Discrepancy	5	4.7	4.0	86,528
Large Household	10	7.2	6.0	131,362
Long Forms				
Count Discrepancy	5	6.3	4.0	38,193
Large Household	10	12.7	10.0	16,130

#### 4.4.9. Telephone interviewer’s use of coverage edit followup instrument

Telephone interviewer debriefings were conducted at two of the call centers after the operation was completed. A series of questions were asked to a mixed group of telephone interviewers and supervisors. Many concerns were raised on a number of issues, including:

- the number and length of the coverage probes
- the requirement to only speak with person one or two
- training did not fully prepare telephone interviewers for live calls
- online frequently asked questions (FAQs) were not complete and were difficult to access
- lack of a clear understanding by the telephone interviewers about why CEFU was being conducted
- difficulty following the rules on verbatim reading of the questions
- respondent sensitivity to some questions, including those about race and Hispanic origin
- inability to redirect a call to a bilingual interviewer
- limitation to continue the interview only with the respondent who had requested a callback

Many of the suggested changes raised by the telephone interviewers may have made their interviewing go smoother. However, because we did not have time to implement and fully test all of these changes, we could not take the chance that the quality of the data we were collecting would be compromised. Simplifying the wording of the questions and coverage probes would have made the interview shorter and less redundant, but would not have aligned with the

residence rules for Census 2000. Allowing the telephone interviewers to stray from our verbatim wording requirements may have allowed more interviews to be completed, but we would have added variability to the questions, and therefore the responses to many of the collected data items in the interview. Still, many of the concerns listed here should be considered by the designers of the survey instrument, of the automated instrument, and as well as the designers of the telephone interviewer training for the CEFU for the next census.

We tracked the use of the knowledge base to specific screens, but we found that it was very infrequently accessed by the telephone interviewers. The knowledge base was designed specifically for the inbound TQA program, not for the CEFU operation. We provided it for the outbound program just in case it would have been of some assistance. The telephone interviewers found it was not very helpful in several situations, including:

- race and Hispanic origin definitions
- multiple residences
- how the data will be used
- why they are being contacted for this followup
- confidentiality requirements

Telephone interviewers also requested a way for the respondent to verify that the call was legitimately from the Census Bureau. They suggested sending advance notification of the call or providing a number for the respondent to call us back.

#### *4.4.10. Collecting person data*

The main objectives of the CEFU operation were to collect person data for large household cases and to correct the roster for count discrepancy cases. Correctness, not coverage gain, was the first and foremost priority. This operation enabled us to correct the respondent-reported household roster of each case by adding or removing names.

Table 36 shows the outcomes of the records for person data that were returned from the CEFU operation. The vast majority of persons listed on rosters of CEFU cases were data defined on the mail back form. None of these 5,664,179 persons had data collected during this operation because there was no content followup component of CEFU in Census 2000.

The CEFU did collect data for 1,388,485 persons who were listed on the continuation rosters of the mail back forms. The demographic data for these people would have been imputed if we had been unable to complete their CEFU interview. In addition, data were collected for 152,683 added persons during CEFU. These added persons would not have been enumerated in Census 2000 if they had not been added through CEFU because they were not included on the mail back rosters. Data were not collected for the 257,882 persons removed from rosters (deletes and duplicates) and would have been enumerated erroneously had they not been corrected in CEFU.

**Table 36. Distribution of person records in completed edit failure cases**

Person records	Edit Failure Type					
	Persons from Completed Forms		Count Discrepancy		Large Household	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Persons</b>	<b>7,463,229</b>	<b>100.0%</b>	<b>1,520,048</b>	<b>100.0%</b>	<b>5,943,181</b>	<b>100.0%</b>
Data Defined Persons	5,664,179	75.9%	1,231,091	81.0%	4,433,088	74.6%
Continuation Roster Persons	1,388,485	18.6%	60,729	4.0%	1,327,756	22.3%
Adds During the CEFU	152,683	2.0%	80,899	5.3%	71,784	1.2%
Deletes During the CEFU	207,182	2.8%	113,845	7.5%	93,337	1.6%
Duplicates During the CEFU	50,700	0.7%	33,484	2.2%	17,216	0.3%

#### 4.4.11. *When were files returned to the Census Bureau from the contractor?*

Overall, there were sixty-two deliveries of evaluation files containing completed cases. These deliveries each contained a file of case level records and another file of person records and were delivered at the same time as the production files. The first delivery was on June 2, 2000 and the final delivery was dated August 16, 2000.

For a table of file delivery sizes by date and number of household and person records, see Table B1 in Appendix B. Note that on July 30, 2000, we received 9,662 household level records but only 1,744 person records. Apparently, an error was made in creating the evaluation files on this day. We should have better testing of file creation software and improve monitoring of files received from contractors in the future.

#### **4.5. How costly were these coverage improvements?**

The contract for CEFU was included in the contract for the TQA program. This program was a short duration program implemented to assist the public in completing their census forms or obtaining information about the census. Since some of the item costs for both the inbound (TQA) and outbound (CEFU) components were not billed separately by the contractor, we were not able to accurately report the separated costs for the CEFU program for these item costs. Moreover, we were not able to report the true value of the total cost of the CEFU operation. In addition, headquarter costs were not included in the cost figures that we do have available.

Therefore, there is little we can discuss of the costs of the coverage edits followup operation. The TQA contract, which includes the cost of the two programs combined, was appropriated at \$102 million. Approximately \$89 million was actually spent on the two programs. The positive variance of \$13 million was the result of lower contractor costs in running the program since the

number of inbound calls of six million was 45 percent lower than the 11 million calls planned (U.S. General Accounting Office, 2001).

Table 37 shows the known itemized costs for the CEFU and TQA operations. Each task shown can be attributed to CEFU only, TQA only, or to both. The tasks on planning and definition, design and development, training, and quality assurance cannot be split out between the two programs. These costs totaled \$56,598,904.83. The TQA costs (inbound) included fulfillment development, fulfillment operations and inbound specific operational costs. These totaled to \$25,533,987.64. The CEFU only (outbound) specific costs were for telephone interviewer’s pay for outbound operations. This totaled \$10,380,182.94. Note that none of these items include headquarters resources or staffing costs.

**Table 37. Cost summary for the coverage edit followup program and the telephone questionnaire assistance program**

Description	Cost
<b>Total Shared Costs</b>	<b>\$56,598,904.83</b>
Planning and Definition	\$1,634,483.75
Design and Development	\$35,223,550.56
Training	\$9,794,959.56
Quality Assurance	\$6,418,592.92
FTS2000* Costs	\$3,527,318.04
<b>Total TQA Only Costs</b>	<b>\$25,533,987.64</b>
TQA Operations	\$24,469,189.06
Fulfillment Development	\$121,168.35
Fulfillment Operations	\$253,753.23
Postage for Fulfillment	\$689,877.00
<b>Total CEFU Only Costs</b>	<b>\$10,380,182.94</b>
CEFU Operations	\$10,380,182.94
<b>Total Costs for CEFU and TQA Combined</b>	<b>\$92,513,075.41</b>

\* Federal telecommunications services

Reporting of costs split by program was not requested during the original contract award. When a detailed cost analysis was later requested, the contractor was unable to change the cost reporting system that was already put in place. In the future, it would be valuable to have better reporting of cost data in order to better understand the true cost of the coverage improvements gained from coverage edit followup.

## 5. RECOMMENDATIONS

### What are the Recommendations?

Given the results and limitations of the data, here are some recommendations for the Coverage Edit Followup operation in Census 2010:

- ▶ Continue to conduct a coverage edit followup operation in future censuses. Include count discrepancy cases and large household cases, as well as other cases we can identify as having a significant possibility of coverage problems.
  
- ▶ Develop ways to increase the completion rate for Coverage Edit Followup operations in the future. We should:
  - Conduct a field followup for cases we do not reach by telephone. This is especially necessary for Puerto Rico and other areas which typically do not have telephones; all cases deserve a followup.
  - Improve our ability to obtain correct telephone numbers for the respondent.
  - Conduct a refusal conversion operation by telephone or field followup to improve the completion rate.
  - Allow interviewers to leave a message when respondents are unavailable so they may call us back to complete the followup.
  
- ▶ Improve case file creation, management, software testing and transmittal procedures of input and output files to avoid loss of data and to ensure information is available to conduct interviews as planned. We should:
  - Improve testing of the universe selection software to avoid selecting ineligible cases for followup and to avoid missing key variables on the input files.
  - Ensure that attempted cases are representative of the entire universe of coverage edit cases in the event the full originally selected universe cannot be followed up.
  - Improve testing and monitoring of files received from contractors in the future to ensure their completeness and accuracy.
  
- ▶ Improve the design of the coverage edit followup instrument to improve effectiveness and reduce respondent burden. We should:
  - Allow telephone interviewers' input into the design of the survey instrument earlier in the development process.
  - Tailor the probe questions to the specific edit failure reason based on the results of this operation during Census 2000 and other relevant research.
  
- ▶ Collect evaluation data in future census tests of coverage followup operations to help improve the methodology used to conduct followup interviews. Ensure we can:

- Collect and analyze the number of call attempts for use in establishing contact with households as well as the number of attempts needed to complete cases in a telephone followup operation.
- Collect and analyze program cost data to better understand the true cost of the coverage improvements gained from coverage edit followup.
  
- ▶ Assign the final household size for count discrepancy cases not completed during coverage edit followup by more closely mimicking the results for completed cases in Census 2000.

## REFERENCES

STSD 1990 REX Memorandum Series # R-3, undated, from James B. Treat, Census Evaluation Branch, Statistical Support Division, to Deborah H. Griffin, Chief, Census Evaluation Branch, Statistical Support Division, Subject: Summary of the 1990 Decennial Census Coverage Edits.

STSD 1990 Decennial Census Memorandum Series # R-1, revision, dated July 22, 1991, from John H. Thompson, Chief, Statistical Support Division, to Arnold A. Jackson, Chief, Decennial Operation Division, and Marvin L. Postma, Acting Chief, Field Division, Subject: Coverage Edit Specifications for the 1990 Decennial Census.

DSSD 2000 Census Memorandum Series # H-13, dated March 2, 1994, from John H. Thompson, Chief, Decennial Statistical Support Division, to Susan M. Miskura, Chief, Year 2000 Research and Development Staff, Subject: Roster Research Results from the Living Situation Survey.

DSSD 1995 Census Test Memorandum Series # E-1; dated June 16, 1994, from John H. Thompson, Chief, Decennial Statistical Support Division, to Susan M. Miskura, Acting Chief, Decennial Management Division, Subject: 1995 Questionnaire Design - Revised Recommended Rostering and Coverage Questions.

Pausche, Joan M., "1994 National Census Test: Quantitative Comparisons of Coverage." Paper presented at the American Statistical Association Annual Meeting in Toronto, Canada. August 15, 1994.

DSSD 1995 Census Test Memorandum Series # E-3; dated September 7, 1994, from John H. Thompson, Chief, Decennial Statistical Support Division, to Susan M. Miskura, Acting Chief, Decennial Management Division, Subject: Evaluation Requirements Document for the 1995 Census Test Research Objective: Coverage Questions for a Complete Listing of Household Members.

DSSD 1995 Census Test Memorandum Series # E-5; dated January 16, 1996, from Ruth Ann Killion, Chief, Decennial Statistical Support Division, to John H. Thompson, Chief, Decennial Management Division, Subject: Coverage Questions Results Evaluation.

1996 Community Census Results Memorandum Series No. 16; dated February 9, 1998, from Ruth Ann Killion, Chief, Decennial Statistical Support Division, to John H. Thompson, Associate Director for Decennial Census, Subject: Results of the Automated Coverage Edits Evaluation.

Census 2000 Dress Rehearsal Evaluation Results Memorandum Series # D-3, dated May 1999, from Ruth Ann Killion, Chief, Planning, Research, and Evaluation Division, and Howard R. Hogan, Decennial Statistical Studies Division, to Kenneth Prewitt, Director, and

William G. Barron, Deputy Director of the Census Bureau, and Paula J. Schneider, Principal Associate Director for Programs, Subject: Coverage Edits Followup Evaluation.

DSSD Census 2000 Procedures and Operations Memorandum Series #DD-3, dated May 18, 1999, from Howard Hogan, Chief, Decennial Statistical Studies Division, to Michael J. Longini, Chief, Decennial Systems and Contracts Management Office, Subject: Specifications for the Identification of Coverage Edit Failures in Census 2000.

Census 2000 DSCMO General Memorandum Series #01-01, dated June 12, 2001, from Michael J. Longini, Chief, Decennial Systems and Contracts Management Office, to Distribution List, Subject: *Telephone Questionnaire Assistance (TQA) and Coverage Edit Follow-up (CEFU) Lessons Learned for Census 2000 - Revised*, DSCMO.

U.S. General Accounting Office, 2001, *Analysis of Fiscal Year 2000 Budget and Internal Control Weaknesses at the U.S. Census Bureau*, Report to Congressional Requesters, December 2001.

U.S. Bureau of the Census, 2002c, *A.1a - Evaluation of the Telephone Questionnaire Assistance Program*, DSSD Census 2000 Evaluations.

U.S. Bureau of the Census, 2002d, *R.1b - Coverage Edit Followup System Requirements Study*, PRED Census 2000 Evaluations.

U.S. Bureau of the Census, 2002e, *Census 2000 Coverage Edit Follow-up Comprehensive Operational Assessment*, Final Draft April 15, 2002, DMD.



## APPENDIX A: Coverage Edit Followup Probe Questions

### Add Probes

- Q1** “Other than those persons you listed, were there any **children** who were living or staying there around the beginning of April? Be sure to consider any newborns, foster children, step children, or children in shared custody arrangements.”
- Q2** “Other than those you listed, were there any **relatives**, such as aunts, uncles, grandparents, cousins, or any other kinds of relatives who were living or staying there around the beginning of April?”
- Q3** “Other than those you listed, were there any **other persons not related to you** who were living or staying there around the beginning of April? For example, someone who rents a room from you or a friend staying with you temporarily while looking for a place to live.”
- Q4** “Other than those you listed, were there **any persons who were either away temporarily or moving** around the beginning of April? For example, a household member who was visiting with friends or relatives, on vacation, on a business trip, or in the process of moving.”
- Q5** “Think back to the beginning of April. Were there **any people staying there who had no other permanent place to stay**? Please tell me their names even if you do not consider them to be regular members of your household”

### Delete Probes

- Q6** “Were any of these people **college students** in April?”
- Q7** “Were any of these people **members of the U. S. Armed Forces** in April?”
- Q8** “Were any of these people **living away in a place such as a prison or jail, mental hospital, nursing home**, or similar place on April 1?”
- Q9** “Some people have more than one place to live. Examples include a second residence where they stay to be closer to work, a friend’s or relative’s home, or a vacation home. Did any people on the list I read you earlier **have another place where they live or stay**?”

**APPENDIX B:**

**Complete edit failure case deliveries**

**Table B1. Daily file delivery size of completed coverage edit followup cases from the contractor by type of record**

Delivery Date	Number of Cases	Number of Person Records
June 2, 2000	368,646	2,069,749
June 6, 2000	29,936	148,717
June 7, 2000	16,483	86,439
June 8, 2000	20,681	119,791
June 9, 2000	19,238	114,532
June 10, 2000	15,458	90,261
June 11, 2000	43,539	237,211
June 13, 2000	28,323	169,021
June 15, 2000	38,334	232,833
June 16, 2000	17,317	108,286
June 17, 2000	13,658	76,682
June 18, 2000	13,862	77,550
June 19, 2000	10,560	58,168
June 20, 2000	16,864	99,612
June 22, 2000	35,952	213,411
June 23, 2000	17,227	90,556
June 24, 2000	14,645	76,912
June 25, 2000	14,741	89,899
June 26, 2000	10,742	54,606
June 27, 2000	18,490	101,433
June 29, 2000	39,776	230,499
June 30, 2000	22,341	138,365
July 1, 2000	11,872	82,690
July 2, 2000	6,143	41,034
July 3, 2000	19,085	107,730
July 4, 2000	20,149	127,765
July 5, 2000	3,685	25,234
July 6, 2000	15,789	103,872
July 7, 2000	17,170	114,625
July 8, 2000	16,182	109,907
July 9, 2000	15,948	103,823
July 10, 2000	12,501	78,579
July 11, 2000	1,662	100,383
July 13, 2000	26,850	184,479
July 14, 2000	10,461	73,161
July 15, 2000	10,178	66,874

Delivery Date	Number of Cases	Number of Person Records
July 16, 2000	13,617	86,067
July 17, 2000	10,002	63,777
July 18, 2000	11,749	81,060
July 20, 2000	22,672	148,136
July 21, 2000	11,708	76,984
July 23, 2000	14,658	91,258
July 24, 2000	5,597	31,713
July 25, 2000	8,120	55,590
July 26, 2000	9,783	62,812
July 27, 2000	9,358	60,463
July 28, 2000	7,424	48,978
July 29, 2000	5,745	38,165
July 30, 2000	9,662	1,743
July 31, 2000	5,412	32,065
August 1, 2000	8,062	53,322
August 3, 2000	14,700	99,096
August 4, 2000	18,280	122,395
August 5, 2000	6,148	39,604
August 6, 2000	6,861	44,487
August 7, 2000	5,477	35,203
August 8, 2000	7,632	50,425
August 10, 2000	11,588	73,640
August 11, 2000	4,212	29,421
August 12, 2000	3,240	21,930
August 13, 2000	3,388	21,948
August 16, 2000	11,757	78,152
<b>Total</b>	<b>1,261,340</b>	<b>7,553,123</b>

**APPENDIX C: More about adds and added persons**

**Table C1. Distribution of completed cases by the number of person adds by count discrepancy type**

Adds Per Case	Count Discrepancy Type					
	Count Discrepancy		High Data Defined Person Count		Low Data Defined Person Count	
	Number	Percent	Number	Percent	Number	Percent
<b>Total number of cases</b>	<b>487,265</b>	<b>100.0%</b>	<b>322,509</b>	<b>100.0%</b>	<b>164,756</b>	<b>100.0%</b>
0 Adds	427,266	87.7%	307,862	95.5%	119,404	72.5%
One or More Adds	59,999	12.3%	14,647	4.5%	45,352	27.5%

Source: CEFU evaluation HH level files - variable *cadd*

**Table C2. Distribution of completed cases by the number of person adds by large household type**

Adds Per Case	Large Household Type					
	Large Household		Large household (more than 6 residents*)		Possible Large Household	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>764,706</b>	<b>100.0%</b>	<b>732,157</b>	<b>100.0%</b>	<b>32,549</b>	<b>100.0%</b>
0 Adds	718,488	94.0%	688,117	94.0%	30,371	93.3%
One or More Adds	46,218	6.0%	44,040	6.0%	2,178	6.7%

Source: CEFU evaluation HH level files - variable *cadd*

**Table C3. Distribution of adds from coverage probe #1 - adding children**

Question 1 - Children		Completed Cases	
		Number	Percent
Complete Cases	<b>Total number of cases</b>	<b>1,251,971</b>	<b>100.0%</b>
	Said no additional children in household <sup>a</sup>	1,219,512	97.4%
	Said there were more children in household <sup>a</sup>	32,459	2.6%
	Did not supply at least one child's name <sup>b</sup>	4,940	15.2%
	Supplied at least one child's name <sup>b</sup>	27,519	84.8%
	Did not add at least one name to the roster <sup>c</sup>	4,903	17.8%
	Added one or more names to the roster <sup>c</sup>	22,616	82.2%
Persons	Names of children added to the household roster	33,246	

a - % is based on 'Total number of cases'

b - % is based on 'Said there were more children in household'

c - % is based on 'Supplied at least one child's name'

Sources: CEFU evaluation HH level files

CEFU evaluation person level files

**Table C4. Number and percent of adds from coverage probe #1 by edit failure reason - adding children**

Question 1 - Children	Edit Failure Reason									
	Completed Cases		Count Discrepancy Cases				Large Household Cases			
			High Data Defined Persons		Low data defined persons		Definite		Possible	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	#	Percent
<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>322,509</b>	<b>25.8%</b>	<b>164,756</b>	<b>13.2%</b>	<b>732,157</b>	<b>58.5%</b>	<b>32,549</b>	<b>2.6%</b>
Said No Additional Children in Household	1,219,512	97.4%	317,809	98.5%	154,274	93.6%	715,650	97.7%	31,779	97.6%
Said There Were More Children in Household	32,459	2.6%	4,700	1.5%	10,482	6.4%	16,507	2.3%	770	2.4%
Did Not Supply at Least One Child's Name	4,940	15.2%	945	20.1%	808	7.7%	3,060	18.5%	127	16.5%
Supplied at Least One Child's Name	27,519	84.8%	3,755	79.9%	9,674	92.3%	13,447	81.5%	643	83.5%
Added 0 Names to the Roster	4,903	17.8%	1,123	29.9%	902	9.3%	2,770	20.6%	108	16.8%
Added 1+ Names to the Roster	22,616	82.2%	2,632	70.1%	8,772	90.7%	10,677	79.4%	535	83.2%
Percent of Cases with Adds		1.81%		0.82%		5.32%		1.46%		1.64%
<b>Names of Children Added to the Household Roster</b>	<b>33,246</b>	<b>100.0%</b>	<b>3,489</b>	<b>10.5%</b>	<b>12,967</b>	<b>39.0%</b>	<b>16,038</b>	<b>48.2%</b>	<b>752</b>	<b>2.3%</b>
Average Number of Person Adds Per Case with Adds	<b>1.47</b>		1.33		1.48		1.50		1.41	

Complete Cases

Persons

**Table C5. Distribution of adds from coverage probe #2 - adding relatives**

Question 2 - Relatives		Completed Cases	
		Number	Percent
Complete Cases	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>
	Said No Additional Relatives in Household <sup>a</sup>	1,234,578	98.6%
	Said There Were More Relatives in Household <sup>a</sup>	17,393	1.4%
	Did Not Supply at Least One Relative's Name <sup>B</sup>	2,138	12.3%
	Supplied at Least One Relative's Name <sup>b</sup>	15,255	87.7%
	Did Not Add at Least One Name to the Roster <sup>c</sup>	1,985	13.0%
	Added One or More Names to the Roster <sup>c</sup>	13,270	87.0%
Persons	Names of Relatives Added to the Household Roster	17,088	

a - % is based on 'Total number of cases'

b - % is based on 'Said there were more relatives in household'

c - % is based on 'Supplied at least one relative's name'

Sources: CEFU evaluation HH level files

CEFU evaluation person level files

**Table C6 - Number and percent of adds from coverage probe #2 by edit failure reason**

		Edit Failure Reason									
		Completed Cases		Count Discrepancy Cases				Large Household Cases			
				High Data Defined Persons		Low data defined persons		Definite		Possible	
Question # 2 - Relatives	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	
<b>Complete Cases</b>	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>322,509</b>	<b>25.8%</b>	<b>164,756</b>	<b>13.2%</b>	<b>732,157</b>	<b>58.5%</b>	<b>32,549</b>	<b>2.6%</b>
	Said No Additional Relatives in Household	1,234,578	98.6%	320,175	99.3%	160,486	97.4%	721,833	98.6%	32,084	98.6%
	Said There Were More Relatives in Household	17,393	1.4%	2,334	0.7%	4,270	2.6%	10,324	1.4%	465	1.4%
	Did Not Supply at Least One Relatives's Name	2,138	12.3%	331	14.2%	266	6.2%	1,469	14.2%	72	15.5%
	Supplied at Least One Relatives's Name	15,255	87.7%	2,003	85.8%	4,004	93.8%	8,855	85.8%	393	84.5%
	Added 0 Names to the Roster	1,985	13.0%	403	20.1%	339	8.5%	1,197	13.5%	46	11.7%
	Added 1+ Names to the Roster	13,270	87.0%	1,600	79.9%	3,665	91.5%	7,658	86.5%	347	88.3%
	Percent of Cases with Adds		1.06%		0.50%		2.22%		1.05%		1.07%
<b>Persons</b>	<b>Names of Relatives Added to the Household Roster</b>	<b>17,088</b>	<b>100.0%</b>	<b>1,842</b>	<b>10.8%</b>	<b>4,277</b>	<b>25.0%</b>	<b>10,532</b>	<b>61.6%</b>	<b>437</b>	<b>2.6%</b>
	Average Number of Person Adds Per Case with Adds	1.29		1.15		1.17		1.38		1.26	



**Table C7. Distribution of adds from coverage probe #3 - adding non relatives**

Question 3 - Adding non Relatives		Completed Cases	
		Number	Percent
Complete Cases	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>
	Said No Additional non Relatives in Household <sup>a</sup>	1,239,673	99.0%
	Said There Were More non Relatives in Household <sup>a</sup>	12,298	1.0%
	Did Not Supply at Least One non Relative's Name <sup>b</sup>	2,046	16.6%
	Supplied at Least One non Relative's Name <sup>b</sup>	10,252	83.4%
	Did Not Add at Least One Name to the Roster <sup>c</sup>	1,687	16.5%
	Added One or More Names to the Roster <sup>c</sup>	8,565	83.5%
Persons	Names of non Relatives Added to the Household Roster	10,442	

a - % is based on 'Total number of cases'

b - % is based on 'Said there were more relatives in household'

c - % is based on 'Supplied at least one relative's name'

Sources: CEFU evaluation HH level files

CEFU evaluation person level files

**Table C8. Number and percent of adds from coverage probe #3 by edit failure reason**

Question # 3 - non Relatives		Edit Failure Reason									
		Completed Cases		Count Discrepancy Cases				Large Household Cases			
				High Data Defined Persons		Low data defined persons		Definite		Possible	
		Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
<b>Complete Cases</b>	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>322,509</b>	<b>25.8%</b>	<b>164,756</b>	<b>13.2%</b>	<b>732,157</b>	<b>58.5%</b>	<b>32,549</b>	<b>2.6%</b>
	Said No Additional non Relatives in Household	1,239,673	99.0%	320,493	99.4%	161,288	97.9%	725,669	99.1%	32,223	99.0%
	Said There Were More non Relatives	12,298	1.0%	2,016	0.6%	3,468	2.1%	6,488	0.9%	326	1.0%
	Did Not Supply at Least One non Relatives's Name	2,046	16.6%	405	20.1%	301	8.7%	1,294	19.9%	46	14.1%
	Supplied at Least One non Relatives's Name	10,252	83.4%	1,611	79.9%	3,167	91.3%	5,194	80.1%	280	85.9%
	Added 0 Names to the Roster	1,687	16.5%	353	21.9%	430	13.6%	859	16.5%	45	16.1%
	Added 1+ Names to the Roster	8,565	83.5%	1,258	78.1%	2,737	86.4%	4,335	83.5%	235	83.9%
	Percent of Cases with Adds		0.68%		0.39%		1.66%		0.59%		0.72%
<b>Persons</b>	<b>Names of non Relatives Added to the Household Roster</b>	<b>10,442</b>	<b>100.0%</b>	<b>1,364</b>	<b>13.1%</b>	<b>3,013</b>	<b>28.9%</b>	<b>5,760</b>	<b>55.2%</b>	<b>305</b>	<b>2.9%</b>
	Average Number of Person Adds per Case with Adds	1.22		1.08		1.10		1.33		1.30	

**Table C9. Distribution of adds from coverage probe #4 - adding persons Moving or temporarily away from the household**

Question 4 - Persons Moving or Temporarily Away		Completed Cases	
		Number	Percent
Complete Cases	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>
	Said No Additional Persons in Household <sup>a</sup>	1,242,789	99.3%
	Said There Were More Persons in Household <sup>a</sup>	9,182	0.7%
	Did Not Supply at Least One Person's Name <sup>B</sup>	3,000	32.7%
	Supplied at Least One Person's Name <sup>b</sup>	6,182	67.3%
	Did Not Add at Least One Name to the Roster <sup>c</sup>	3,029	49.0%
Added One or More Names to the Roster <sup>c</sup>	3,153	51.0%	
Persons	Names of Persons Added to the Household Roster	3,440	

a - % is based on 'Total number of cases'

b - % is based on 'Said there were more relatives in household'

c - % is based on 'Supplied at least one relative's name'

Sources: CEFU evaluation HH level files

CEFU evaluation person level files

**Table C10. Number and Percent of adds from coverage probe # 4 - persons temporarily away- by edit failure reason**

Question #	Edit Failure Reason									
	Completed Cases		Count Discrepancy Cases				Large Household Cases			
			High Data Defined Persons		Low Data Defined Persons		Definite		Possible	
#	%	#	%	#	%	#	%	#	%	
<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>322,509</b>	<b>25.8%</b>	<b>164,756</b>	<b>13.2%</b>	<b>732,157</b>	<b>58.5%</b>	<b>32,549</b>	<b>2.6%</b>
Said No Persons Temp Away from Household	1,242,789	99.3%	320,870	99.5%	163,533	99.3%	726,125	99.2%	32,261	99.1%
Said There Were More Persons Temp Away from Household	9,182	0.7%	1,639	0.5%	1,223	0.7%	6,032	0.8%	288	0.9%
Did Not Supply the Name of at Least One Temp Away Person Name	3,000	32.7%	678	41.4%	211	17.3%	2,019	33.5%	92	31.9%
Supplied at Least One Name of a Person Temp Away	6,182	67.3%	961	58.6%	1,012	82.7%	4,013	66.5%	196	68.1%
Added 0 Names to the Roster	3,029	49.0%	568	59.1%	514	50.8%	1,862	46.4%	85	43.4%
Added 1+ Names to the Roster	3,153	51.0%	393	40.9%	498	49.2%	2,151	53.6%	111	56.6%
Percent of Cases with Adds		0.25%		0.12%		0.30%		0.29%		0.34%
<b>Names of Persons Temp Away from Household Added to the Roster</b>	<b>3,440</b>	<b>100.0%</b>	<b>381</b>	<b>11.1%</b>	<b>513</b>	<b>14.9%</b>	<b>2,422</b>	<b>70.4%</b>	<b>124</b>	<b>3.6%</b>
Average Number of Person Adds per Case with Adds	1.14		0.67		1.00		1.30		1.46	

Complete Cases

Persons

**Table C11. Distribution of adds from coverage probe #5 - adding persons with no other permanent place to live**

Question 1 - Persons with No Other Permanent Place to Live		Completed Cases	
		Number	Percent
Complete Cases	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>
	Said No Additional Persons in Household <sup>a</sup>	1,243,392	99.3%
	Said There Were More Persons in Household <sup>a</sup>	8,579	0.7%
	Did Not Supply at Least One Person's Name <sup>B</sup>	3,121	36.4%
	Supplied at Least One Person's Name <sup>b</sup>	5,458	63.6%
	Did Not Add at Least One Name to the Roster <sup>c</sup>	1,782	32.6%
	Added One or More Names to the Roster <sup>c</sup>	3,676	67.4%
Persons	Names of Persons Added to the Household Roster	4,436	

a - % is based on 'Total number of cases'

b - % is based on 'Said there were more relatives in household'

c - % is based on 'Supplied at least one relative's name'

Sources: CEFU evaluation HH level files

CEFU evaluation person level files

**Table C12. Number and percent of adds from coverage probe #5 by edit failure reason**

Question # 5 - Persons with No Permanent Place to Stay		Edit Failure Reason									
		Completed Cases		Count Discrepancy Cases				Large Household Cases			
				High Data Defined Persons		Low Data Defined Persons		Definite		Possible	
Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent		
<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>322,509</b>	<b>25.8%</b>	<b>164,756</b>	<b>13.2%</b>	<b>732,157</b>	<b>58.5%</b>	<b>32,549</b>	<b>2.6%</b>	
Complete Cases	Said No Persons Without a Permanent Place to Stay in Household	1,243,392	99.3%	321,215	99.6%	163,548	99.3%	726,350	99.2%	32,279	99.2%
	Said There Were Persons Without a Permanent Place to Stay	8,579	0.7%	1,294	0.4%	1,208	0.7%	5,807	0.8%	270	0.8%
	Did Not Supply the Name of Any Persons Without a Permanent Place to Stay	3,121	36.4%	422	32.6%	207	17.1%	2,390	41.2%	102	37.8%
	Supplied at Least One Name of a Person Without a Permanent Place to Stay	5,458	63.6%	872	67.4%	1,001	82.9%	3,417	58.8%	168	62.2%
	Added No Names to the Roster	1,782	32.6%	322	36.9%	251	25.1%	1,158	33.9%	51	30.4%
	Added 1+ Names to the Roster	3,676	67.4%	550	63.1%	750	74.9%	2,259	66.1%	117	69.6%
	Percent of Cases with Adds		0.29%		0.17%		0.46%		0.31%		0.36%
Persons	<b>Names of Person Without a Permanent Place to Stay Added to the Household Roster</b>	<b>4,436</b>	<b>100.0%</b>	<b>621</b>	<b>14.0%</b>	<b>840</b>	<b>18.9%</b>	<b>2,829</b>	<b>63.8%</b>	<b>146</b>	<b>3.3%</b>
	Average Number of Person Adds Per Case with Adds	1.21		1.13		1.12		1.25		1.25	

**Table C13. Distribution of adds from interrupt option - adding persons**

Question 1 - Persons Added Through the Interrupt Option		Completed Cases	
		Number	Percent
Complete Cases	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>
	No Interruption to Add Additional Persons <sup>a</sup>	1,187,270	94.8%
	Interruption to Add Additional Persons <sup>a</sup>	64,701	5.2%
	Did Not Add at Least One Name to the Roster <sup>b</sup>	1,510	2.3%
	Added One or More Names to the Roster <sup>b</sup>	63,191	97.7%
Persons	Names of Persons Added to the Household Roster	83,497	

a - % is based on 'Total number of cases'

b - % is based on 'Interruption to add additional persons'

Sources: CEFU evaluation HH level files

CEFU evaluation person level files

**Table C14. Number and percent of adds from interruption option by edit failure reason**

Question #		Edit Failure Reason									
		Count Discrepancy Cases						Large Household Cases			
		Completed Cases		High Data Defined Persons		Low Data Defined Persons		Definite		Possible	
		Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
<b>Complete Cases</b>	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>322,509</b>	<b>25.8%</b>	<b>164,756</b>	<b>13.2%</b>	<b>732,157</b>	<b>58.5%</b>	<b>32,549</b>	<b>2.6%</b>
	Did Not Try an Interrupt Add	1,187,270	94.8%	312,822	97.0%	132,848	80.6%	710,098	97.0%	31,502	96.8%
	Did Try an Interrupt Add	64,701	5.2%	9,687	3.0%	31,908	19.4%	22,059	3.0%	1,047	3.2%
	Added 0 Names to the Roster	1,510	2.3%	437	4.5%	363	1.1%	671	3.0%	39	3.7%
	Added 1+ Names to the Roster	63,191	97.7%	9,250	95.5%	31,545	98.9%	21,388	97.0%	1,008	96.3%
	Percent of Cases with Adds		5.05%		2.87%		19.15%		2.92%		3.10%
<b>Persons</b>	<b>Names of Children Added to the Household Roster</b>	<b>83,497</b>	<b>100.0%</b>	<b>11,152</b>	<b>13.4%</b>	<b>40,192</b>	<b>48.1%</b>	<b>30,852</b>	<b>37.0%</b>	<b>1,301</b>	<b>1.6%</b>
	Average Number of Person Adds Per Case with Adds	1.32		1.21		1.27		1.44		1.29	



## Appendix D: More about deletes and deleted persons

**Table D1. Distribution of deleted persons per completed edit failure case by count discrepancy type**

Deletes per Case	Count Discrepancy Type					
	Count Discrepancy		High Data Defined Person Count		Low Data Defined Person Count	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>487,265</b>	<b>100.0%</b>	<b>322,509</b>	<b>100.0%</b>	<b>164,756</b>	<b>100.0%</b>
Cases with No Deletes	409,517	84.0%	260,495	80.8%	149,022	90.5%
Cases with One or More Deletes	77,748	16.0%	62,014	19.2%	15,734	9.6%

Source: CEFU evaluation HH level files - variable cdel

**Table D2. Distribution of deleted persons per completed edit failure case by large household type**

Adds Per Case	Large Household Type					
	Large Household		Large Household (More than 6 Residents*)		Possible Large Household	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>764,706</b>	<b>100.0%</b>	<b>732,157</b>	<b>100.0%</b>	<b>32,549</b>	<b>100.0%</b>
Cases with No Deletes	715,333	93.5%	685,375	93.6%	29,958	92.0%
Cases with One or More Deletes	49,373	6.5%	46,782	6.4%	2,591	8.0%

Source: CEFU evaluation HH level files - variable cdel

**Table D3. Distribution of deletes from probe question #6 - college students**

Question 1 - Relatives		Completed Cases	
		#	%
Complete Cases	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>
	Said No College Students in Household <sup>a</sup>	1,217,886	97.3%
	Said There Were College Students in Household <sup>a</sup>	34,085	2.7%
	Did Not Supply at Least College Student's Name <sup>B</sup>	6,402	18.8%
	Supplied at Least One College Student's Name <sup>B</sup>	27,683	81.2%
	Did Not Delete at Least One Name from the Roster <sup>C</sup>	11,120	40.2%
	Deleted One or More Names from the Roster <sup>C</sup>	16,563	59.8%
Persons	Names of Students Deleted from the Household Roster	19,103	

a - % is based on 'Total number of cases'

b - % is based on Said there were college students in household'

c - % is based on Supplied at least one college student's name'

Sources: CEFU evaluation HH level files

CEFU evaluation person level files

**Table D4 - Number and percent of deletes from coverage probe #6 (college students) by edit failure reason**

Question #	Edit Failure Reason									
	Completed Cases		Count Discrepancy Cases				Large Household Cases			
			High Data Defined Persons		Low Data Defined Persons		Definite		Possible	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>322,509</b>	<b>25.8%</b>	<b>164,756</b>	<b>13.2%</b>	<b>732,157</b>	<b>58.5%</b>	<b>32,549</b>	<b>2.6%</b>
Said No College Students in Household	1,217,886	97.3%	310,654	96.3%	161,816	98.2%	713,799	97.5%	31,617	97.1%
Said There Were College Students in Household	34,085	2.7%	11,855	3.7%	2,940	1.8%	18,358	2.5%	932	2.9%
Did Not Supply at Least One College Students's Name	6,402	18.8%	1,738	14.7%	880	29.9%	3,612	19.7%	172	18.5%
Supplied at Least One College Students's Name	27,683	81.2%	10,117	85.3%	2,060	70.1%	14,746	80.3%	760	81.5%
Deleted 0 Names from the Roster	11,120	40.2%	2,730	27.0%	1,004	48.7%	7,096	48.1%	290	38.2%
Deleted 1+ Names from the Roster	16,563	59.8%	7,387	73.0%	1,056	51.3%	7,650	51.9%	470	61.8%
Percent of Cases with Deletes		1.32%		2.29%		0.64%		1.04%		1.44%
<b>Names of College Students Deleted to the Household Roster</b>	<b>19,103</b>	<b>100.0%</b>	<b>8,129</b>	<b>42.6%</b>	<b>1,167</b>	<b>6.1%</b>	<b>9,237</b>	<b>48.4%</b>	<b>570</b>	<b>3.0%</b>
Average Number of Person Deletes Per Case with Deletes	1.15		1.10		1.11		1.21		1.21	

Complete Cases

Persons

**Table D5. Distribution of deletes from probe question #7 - military**

		Completed Cases	
		#	%
Question 1 - Relatives			
Complete Cases	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>
	Said No Military Members in Household <sup>a</sup>	1,246,321	99.5%
	Said There Were Military Members in Household <sup>a</sup>	5,650	0.5%
	Did Not Supply at Least Military Members' Name <sup>b</sup>	1,515	26.8%
	Supplied at Least One Military Members' Name <sup>b</sup>	4,135	73.2%
	Did Not Delete at Least One Name from the Roster <sup>c</sup>	2,239	54.1%
	Deleted One or More Names from the Roster <sup>c</sup>	1,896	45.9%
Persons	Names of Military Members Deleted from the Household Roster	2,022	

a - % is based on 'Total number of cases'

b - % is based on Said there were military members in household'

c - % is based on Supplied at least one military members' name'

Sources: CEFU evaluation HH level files

CEFU evaluation person level files

**Table D6. Number and percent of deletes from coverage probe #7 (military) by edit failure reason**

Question #	Edit Failure Reason										
	Completed Cases		Count Discrepancy Cases				Large Household Cases				
			High Data Defined Persons		Low Data Defined Persons		Definite		Possible		
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	
<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>322,509</b>	<b>25.8%</b>	<b>164,756</b>	<b>13.2%</b>	<b>732,157</b>	<b>58.5%</b>	<b>32,549</b>	<b>2.6%</b>	
Said No Military in Household	1,246,321	99.5%	320,941	99.5%	164,128	99.6%	728,842	99.5%	32,410	99.6%	
Said There Were Military in Household	5,650	0.5%	1,568	0.5%	628	0.4%	3,315	0.5%	139	0.4%	
Complete Cases	Did Not Supply at Least One Military Member's Name	1,515	26.8%	350	22.3%	205	32.6%	919	27.7%	41	29.5%
	Supplied at Least One Military Member's Name	4,135	73.2%	1,218	77.7%	423	67.4%	2,396	72.3%	98	70.5%
	Deleted 0 Names from the Roster	2,239	54.1%	533	43.8%	265	62.6%	1,395	58.2%	46	46.9%
	Deleted 1+ Names from the Roster	1,896	45.9%	685	56.2%	158	37.4%	1,001	41.8%	52	53.1%
	Percent of Cases with Deletes		0.15%		0.21%		0.10%		0.14%		0.16%
Persons	<b>Names of Military Members Added to the Household Roster</b>	<b>2,022</b>	<b>100.0%</b>	<b>715</b>	<b>35.4%</b>	<b>169</b>	<b>8.4%</b>	<b>1,085</b>	<b>53.7%</b>	<b>53</b>	<b>2.6%</b>
	Average Number of Person Deletes Per Case with Deletes	1.07		1.04		1.07		1.08		1.02	

**Table D7. Distribution of deletes from probe question #8 - persons in institutions**

Question 1 - Relatives		Completed Cases	
		Number	Percent
Complete Cases	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>
	Said No Persons in Institutions in Household <sup>a</sup>	1,244,318	99.4%
	Said There Were Persons in Institutions in Household <sup>a</sup>	7,653	0.6%
	Did Not Supply at Least One Person in Institutions Name <sup>B</sup>	1,777	23.2%
	Supplied at Least One Persons in Institutions Name <sup>B</sup>	5,876	76.8%
	Did Not Delete at Least One Name from the Roster <sup>C</sup>	1,070	18.2%
	Deleted One or More Names from the Roster <sup>C</sup>	4,806	81.8%
Persons	Names of Persons in Institutions Deleted from the Household Roster	5,165	

a - % is based on 'Total number of cases'

b - % is based on Said there were persons in institutions in household'

c - % is based on Supplied at least one persons in institutions' name'

Sources: CEFU evaluation HH level files

CEFU evaluation person level files

**Table D8. Number and percent of deletes from coverage probe #8 (institutions) by edit failure reason**

Question # 8 - Institutions	Edit Failure Reason									
	Completed Cases		Count Discrepancy Cases				Large Household Cases			
			High Data Defined Persons		Low Data Defined Persons		Definite		Possible	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>322,509</b>	<b>25.8%</b>	<b>164,756</b>	<b>13.2%</b>	<b>732,157</b>	<b>58.5%</b>	<b>32,549</b>	<b>2.6%</b>
Said No Additional Persons in Institutions in Household	1,244,318	99.4%	319,770	99.2%	164,106	99.6%	728,086	99.4%	32,356	99.4%
Said There Were More Persons in Institutions in Household	7,653	0.6%	2,739	0.8%	650	0.4%	4,071	0.6%	193	0.6%
Did Not Supply the Name of at Least One Person in an Institution	1,777	23.2%	520	19.0%	229	35.2%	981	24.1%	47	24.4%
Supplied the Name of at Least One Person in an Institution	5,876	76.8%	2,219	81.0%	421	64.8%	3,090	75.9%	146	75.6%
Deleted 0 Names from the Roster	1,070	18.2%	352	15.9%	86	20.4%	598	19.4%	34	23.3%
Deleted 1+ Names from the Roster	4,806	81.8%	1,867	84.1%	335	79.6%	2,492	80.6%	112	76.7%
Percent of Cases with Deletes		0.38%		0.58%		0.20%		0.34%		0.34%
<b>Names of Persons in Institutions Added to the Household Roster</b>	<b>5,165</b>	<b>100.0%</b>	<b>1,921</b>	<b>37.2%</b>	<b>345</b>	<b>6.7%</b>	<b>2,780</b>	<b>53.8%</b>	<b>119</b>	<b>2.3%</b>
Average Number of Person Deletes Per Case with Deletes	1.07		1.03		1.03		1.12		1.06	

Complete Cases

Persons

**Table D9. Distribution of deletes from probe question #9 - persons with a second residence**

Question 9 - Persons with a Second Residence		Completed Cases	
		Number	Percent
Complete Cases	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>
	Said No Persons with “Another Place Where They Live or Stay” in Household <sup>a</sup>	1,216,569	97.2%
	Said There Were Persons with a “Another Place Where They Live or Stay” in Household <sup>a</sup>	35,402	2.8%
	Did Not Supply the Name of at Least One Person with “Another Place Where They Live or Stay” <sup>b</sup>	6,628	18.7%
	Supplied the Name of at Least One Persons with “Another Place Where They Live or Stay” <sup>b</sup>	28,774	81.3%
	Did Not Delete at Least One Name from the Roster <sup>c</sup>	17,056	59.3%
	Deleted One or More Names from the Roster <sup>c</sup>	11,718	40.7%
Persons	Names of Persons with “Another Place Where They Live or Stay” Deleted from the Household Roster	16,255	

a - % is based on ‘Total number of cases’

b - % is based on ‘Said there were persons with a “another place where they live or stay” in household’

c - % is based on ‘Supplied the name of at least one persons with “another place where they live or stay” ’

Sources: CEFU evaluation HH level files

CEFU evaluation person level files



**Table D10. Number and percent of deletes from coverage probe #9 (second residence) by edit failure reason**

Question # 9 - Second Residence	Edit Failure Reason									
	Completed Cases		Count Discrepancy Cases				Large Household Cases			
			High Data Defined Persons		Low Data Defined Persons		Definite		Possible	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>322,509</b>	<b>25.8%</b>	<b>164,756</b>	<b>13.2%</b>	<b>732,157</b>	<b>58.5%</b>	<b>32,549</b>	<b>2.6%</b>
Said No Additional Person with a Second Residence in Household	1,216,569	97.2%	310,557	96.3%	161,421	98.0%	713,008	97.4%	31,583	97.0%
Said There Were Persons with a Second Residence in Household	35,402	2.8%	11,952	3.7%	3,335	2.0%	19,149	2.6%	966	3.0%
Did Not Supply the Name of a Person with a Second Residence	6,628	18.7%	2,453	20.5%	784	23.5%	3,233	16.9%	158	16.4%
Supplied at Least One Name of a Person with a Second Residence	28,774	81.3%	9,499	79.5%	2,551	76.5%	15,916	83.1%	808	83.6%
Deleted 0 Names from the Roster	17,056	59.3%	4,804	50.6%	1,669	65.4%	10,088	63.4%	495	61.3%
Deleted 1+ Names from the Roster	11,718	40.7%	4,695	49.4%	882	34.6%	5,828	36.6%	313	38.7%
Percent of Cases with Deletes		0.94%		1.46%		0.54%		0.80%		0.96%
<b>Names of Children Added to the Household Roster</b>	<b>16,255</b>	<b>100.0%</b>	<b>6,001</b>	<b>36.9%</b>	<b>1,095</b>	<b>6.7%</b>	<b>8,714</b>	<b>53.6%</b>	<b>445</b>	<b>2.7%</b>
Average Number of Person Deletes Per Case with Deletes	1.39		1.28		1.24		1.50		1.42	

Complete Cases

Persons

**Table D11. Distribution of deletes from interrupt option**

		Completed Cases	
		Number	Percent
Interrupt Option			
Complete Cases	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>
	Did Not Interrupt to Delete a Roster Name <sup>a</sup>	1,149,820	91.8%
	Did Interrupt to Delete a Roster Name <sup>a</sup>	102,151	8.2%
	Did Not Delete at Least One Name from the Roster <sup>B</sup>	7,599	7.4%
	Deleted One or More Names from the Roster <sup>B</sup>	94,552	92.6%
Persons	Names of Persons Deleted from the Household Roster by Interrupting	164,368	

a - % is based on 'Total number of cases'

b - % is based on 'Did interrupt to delete a roster name '

Sources: CEFU evaluation HH level files

    CEFU evaluation person level files

**Table D12. Number and percent of deletes from interruption option by edit failure reason**

Interrupt Deletes		Edit Failure Reason									
		Count Discrepancy Cases						Large Household Cases			
		Completed Cases		High Data Defined Persons		Low Data Defined Persons		Definite		Possible	
		Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>		<b>1,251,971</b>	<b>100.0%</b>	<b>322,509</b>	<b>25.8%</b>	<b>164,756</b>	<b>13.2%</b>	<b>732,157</b>	<b>58.5%</b>	<b>32,549</b>	<b>2.6%</b>
Complete Cases	Did Not Try an Interrupt Delete	1,149,820	91.8%	271,732	84.3%	149,933	91.0%	697,522	95.3%	30,633	94.1%
	Did Try an Interrupt Delete	102,151	8.2%	50,777	15.7%	14,823	9.0%	34,635	4.7%	1,916	5.9%
	Deleted 0 Names from the Roster	7,599	7.4%	2,523	5.0%	1,303	8.8%	3,568	10.3%	205	10.7%
	Deleted 1+ Names from the Roster	94,552	92.6%	48,254	95.0%	13,520	91.2%	31,067	89.7%	1,711	89.3%
	Percent of Cases with Deletes		7.55%		14.96%		8.21%		4.24%		5.26%
Persons	<b>Names of Children Deleted from the Household Roster</b>	<b>164,368</b>	<b>100.0%</b>	<b>74,933</b>	<b>45.6%</b>	<b>19,228</b>	<b>11.7%</b>	<b>65,666</b>	<b>40.0%</b>	<b>4,541</b>	<b>2.8%</b>
	Average Number of Person Deletes per Case with Deletes	1.74		1.55		1.42		2.11		2.65	

Appendix H: Duplicates by coverage probe

**APPENDIX E:**

**Distribution of duplicated persons by CEFU reason**

**Table E1. Distribution of completed cases by the number of person duplicates by count discrepancy type**

Duplicates Per Case	Count Discrepancy Type					
	Count Discrepancy		High Data Defined Person Count		Low Data Defined Person Count	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>487,265</b>	<b>100.0%</b>	<b>322,509</b>	<b>100.0%</b>	<b>164,756</b>	<b>100.0%</b>
Cases with No Duplicates	466,308	95.7%	304,527	94.4%	161,781	98.2%
Cases with One or More Duplicates	20,957	4.3%	17,982	5.6%	2,975	1.8%

Source: CEFU evaluation HH level files - variable cdup

**Table E2. Distribution of completed cases by the number of person duplicates by large household type**

Adds Per Case	Large Household Type					
	Large Household		Large Household (More than 6 Residents*)		Possible Large Household	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>764,706</b>	<b>100.0%</b>	<b>732,157</b>	<b>100.0%</b>	<b>32,549</b>	<b>100.0%</b>
Cases with No Duplicates	758,341	99.2%	726,356	99.2%	31,985	98.3%
Cases with One or More Duplicates	6,365	0.8%	5,801	0.8%	564	1.7%

Source: CEFU evaluation HH level files - variable cdup

**Table E3. Percent of complete cases with names removed from the roster after the last coverage probe because more than one roster name represented the same person**

		Completed Cases	
		Number	Percent
	Last Check Screen		
Complete Cases	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>
	Did Not Delete a Roster Name <sup>a</sup>	1,229,502	98.2%
	Did Delete a Roster Name <sup>a</sup>	22,469	1.8%
Persons	Names of Persons Deleted from the Household Roster by Interrupting	39,419	

a - % is based on 'Total number of cases'

Sources: CEFU evaluation HH level files

CEFU evaluation person level files

**Table E4. Percent of complete cases with names removed from the roster after the last coverage probe because more than one roster name represented the same person by edit failure reason**

Question #		Edit Failure Reason									
		Completed Cases		Count Discrepancy Cases				Large Household Cases			
				High Data Defined Persons		Low Data Defined Persons		Definite		Possible	
		Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
<b>Complete Cases</b>	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>322,509</b>	<b>25.8%</b>	<b>164,756</b>	<b>13.2%</b>	<b>732,157</b>	<b>58.5%</b>	<b>32,549</b>	<b>2.6%</b>
	Deleted 0 Names from the Roster	1,229,502	98.2%	307,081	95.2%	162,430	98.6%	727,872	99.4%	32,119	98.7%
	Deleted 1+ Names from the Roster	22,469	1.8%	15,428	4.8%	2,326	1.4%	4,285	0.6%	430	1.3%
<b>Persons</b>	<b>Number of Names Removed from the Household Roster</b>	<b>39,419</b>	<b>100.0%</b>	<b>23,904</b>	<b>60.6%</b>	<b>3,301</b>	<b>8.4%</b>	<b>10,887</b>	<b>27.6%</b>	<b>1,327</b>	<b>3.4%</b>

**Table E5. Percent of complete cases with names removed from the roster through the interrupt option because more than one roster name represented the Same person**

		Completed Cases	
		Number	Percent
Last Check Screen			
Complete Cases	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>
	Did Not Delete a Roster Name <sup>a</sup>	1,247,037	99.6%
	Did Delete a Roster Name <sup>a</sup>	4,934	0.4%
Persons	<b>Names of Persons Deleted from the Household Roster by Interrupting</b>	<b>11,235</b>	

a - % is based on 'Total number of cases'

Sources: CEFU evaluation HH level files

CEFU evaluation person level files

**Table E6. Number and percent of duplicates from interruption option by edit failure reason**

Question#		Edit Failure Reason									
		Completed Cases		Count Discrepancy Cases				Large Household Cases			
				High Data Defined Persons		Low Data Defined Persons		Definite		Possible	
		Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
<b>Complete Cases</b>	<b>Total Number of Cases</b>	<b>1,251,971</b>	<b>100.0%</b>	<b>322,509</b>	<b>25.8%</b>	<b>164,756</b>	<b>13.2%</b>	<b>732,157</b>	<b>58.5%</b>	<b>32,549</b>	<b>2.6%</b>
	Deleted 0 Names from the Roster	1,247,037	99.6%	319,911	99.2%	164,099	99.6%	730,614	99.8%	32,413	99.6%
	Deleted 1+ Names from the Roster	4,934	0.4%	2,598	0.8%	657	0.4%	1,543	0.2%	136	0.4%
<b>Persons</b>	<b>Number of Names Removed from the Household Roster</b>	<b>11,235</b>	<b>100.0%</b>	<b>5,194</b>	<b>46.2%</b>	<b>1,065</b>	<b>9.5%</b>	<b>4,524</b>	<b>40.3%</b>	<b>452</b>	<b>4.0%</b>



**APPENDIX F: More about name edits**

**Table F1. Distribution of cases with name edits per completed edit failure case by count discrepancy type**

Name Edits per Case	Count Discrepancy Type					
	Count Discrepancy		High Data Defined Person Count		Low Data Defined Person Count	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>487,265</b>	<b>100.0%</b>	<b>322,509</b>	<b>100.0%</b>	<b>164,756</b>	<b>100.0%</b>
Cases with No Name Edits	362,030	74.3%	242,475	75.2%	119,555	72.6%
Cases with One or More Name Edits	125,235	25.7%	80,034	24.8%	45,201	27.4%

Source: CEFU Evaluation Hh Level Files - Variable Cedit

**Table F2. Distribution of cases with name edits per completed edit failure case by large household type**

Name Edits per Case	Large Household Type					
	Large Household		Large Household (More than 6 Residents*)		Possible Large Household	
	Number	Percent	Number	Percent	Number	Percent
<b>Total Number of Cases</b>	<b>764,706</b>	<b>100.0%</b>	<b>732,157</b>	<b>100.0%</b>	<b>32,549</b>	<b>100.0%</b>
Cases with No Name Edits	516,120	67.5%	493,845	67.5%	22,275	68.4%
Cases with One or More Name Edits	248,586	32.5%	238,312	32.5%	10,274	31.6%

Source: CEFU evaluation HH level files - variable cedit

**APPENDIX G: Demographics of persons added, deleted, or marked as duplicates**

**Table G1. Frequency and percent of tenure of housing units with changes in the household roster in coverage edit followup in Census 2000**

Tenure	Cases with CEFU Adds		Cases with CEFU Deletes or Duplicates		All Census Returns	
	Number	Percent	Number	Percent	Number	Percent
<b>Total</b>	<b>106,217</b>	<b>-</b>	<b>151,713</b>	<b>-</b>	<b>106,741,426</b>	<b>-</b>
Owner	71,175	67.7%	100,470	68.2%	70,735,522	66.3%
Renter	33,882	32.3%	46,887	31.8%	36,005,904	33.7%
Missing or Invalid	1,160	-	4,356	-	-	-

Source: Drf2 percents are calculated excluding missing or invalid values

**Table G2. Frequency and percent of sex of persons added or removed from household rosters by coverage edit followup compared to the overall population**

Sex	CEFU Adds		CEFU Deletes and Duplicates		All Census Returns	
	Number	Percent	Number	Percent	Number	Percent
<b>Total</b>	<b>152,683</b>	<b>-</b>	<b>257,882</b>	<b>-</b>	<b>285,230,516</b>	<b>-</b>
Male	76,125	50.9%	63,425	48.3%	139,887,140	49.0%
Female	73,400	49.1%	67,883	51.7%	145,343,376	51.0%
Missing	3,158	-	126,574	-	-	-

Source: DRF2 Percents are calculated excluding Missing values

**Table G3. Frequency and percent of age groups of persons added or removed from household rosters by coverage edit followup compared to the overall population**

Age	CEFU Adds		CEFU Deletes and Duplicates		All Census Returns	
	Number	Percent	Number	Percent	Number	Percent
<b>Total</b>	<b>152,683</b>	<b>-</b>	<b>257,882</b>	<b>-</b>	<b>285,230,516</b>	<b>-</b>
Under 5	13,081	12.0%	3,208	3.3%	19,471,204	6.8%
5 to 9 years	13,242	12.1%	4,485	4.7%	20,854,667	7.3%
10 to 14 years	10,607	9.7%	4,406	4.6%	20,833,872	7.3%
15 to 19 years	11,185	10.2%	13,274	13.8%	20,533,326	7.2%
20 to 24 years	9,726	8.9%	13,350	13.9%	19,265,192	6.8%
25 to 34 years	11,750	10.8%	10,062	10.5%	40,426,056	14.2%
35 to 44 years	9,532	8.7%	8,886	9.2%	45,664,190	16.0%
45 to 54 years	7,551	6.9%	7,810	8.1%	38,140,988	13.4%
55 to 59 years	3,340	3.1%	3,410	3.5%	13,658,120	4.8%
60 to 64 years	3,386	3.1%	3,252	3.4%	10,966,011	3.8%
65 to 74 years	7,479	6.8%	6,995	7.3%	18,631,937	6.5%
75 to 84 years	6,442	5.9%	7,998	8.3%	12,497,660	4.4%
85 years and over	1,878	1.7%	6,073	6.3%	4,287,293	1.5%
Missing or Invalid	43,484	-	161,673	-	-	-

Source: DRF2 Percents are calculated excluding Missing or Invalid values

**Table G4. Frequency and percent of race of persons added or removed from household rosters by coverage edit followup compared to the overall population**

Race	CEFU adds		CEFU deletes and duplicates		All Census Returns	
	Number	Percent	Number	Percent	Number	Percent
<b>Total</b>	<b>152,683</b>	<b>-</b>	<b>257,882</b>	<b>-</b>	<b>285,230,516</b>	<b>-</b>
White	69,568	49.5%	95,801	70.4%	214,525,488	75.2%
Black, African American	28,605	20.4%	33,196	24.4%	34,961,123	12.3%
American Indian, Alaska Native	1,330	0.9%	1,329	1.0%	2,489,292	0.9%
Asian	5,791	4.1%	2,516	1.8%	10,250,958	3.6%
Native Hawaiian or Other Pacific Islander	596	0.4%	119	0.1%	399,928	0.1%
Some Other Race	30,662	21.8%	617	0.5%	15,619,084	5.5%
Two or More	3,976	2.8%	2,531	1.9%	6,984,643	2.4%
Missing	12,155	-	121,773	-	-	-

Source: DRF2 Percents are calculated excluding Missing values

**Table G5. Frequency and percent of Hispanic origin of persons added or removed from household rosters by coverage edit followup compared to the overall population**

Hispanic Origin	CEFU Adds		CEFU Deletes and Duplicates		All Census Returns	
	Number	Percent	Number	Percent	Number	Percent
<b>Total</b>	<b>152,683</b>	<b>-</b>	<b>257,882</b>	<b>-</b>	<b>285,230,516</b>	<b>-</b>
Non-Spanish/Hispanic	93,547	65.7%	107,394	82.7%	246,161,952	86.3%
Mexican, Mexican American, Chicano	35,672	25.1%	10,592	8.2%	20,652,257	7.2%
Puerto Rican	3,591	2.5%	3,515	2.7%	7,029,570	2.5%
Cuban	586	0.4%	981	0.8%	1,261,658	0.4%
Other Spanish/Hispanic	8,987	6.3%	5,556	4.3%	10,125,079	3.6%
Two or More Origin Responses	5	0.0%	1,832	1.4%		
Missing	10,295	-	128,012	-	-	-

Source: DRF2 Percents are calculated excluding Missing values

**APPENDIX H: Demographics of Large Household Persons**

**Table H1. Frequency and Percent of Tenure of Housing Units for Large Household Cases in Coverage Edit Followup by Completion Status**

Tenure	Complete LHHFU Cases		Incomplete LHHFU Cases		All Census 2000 Persons	
	Number	Percent	Number	Percent	Number	Percent
<b>Total</b>	<b>764,706</b>	<b>-</b>	<b>568,072</b>		<b>106,741,426</b>	<b>-</b>
Owner	530,782	70.0%	338,052	60.1%	70,735,522	66.3%
Renter	227,528	30.0%	224,748	39.9%	36,005,904	33.7%
Missing or Invalid	6,396	-	5,272		-	-

Source: DRF2

Percents are calculated excluding Missing or Invalid values

**Table H2. Frequency and Percent of Sex of Persons on Completed Coverage Edit Followup Large Household Cases Compared to the Overall Population by Persons 1-6 and Persons 7-12.**

Sex	Persons from LHHFU Continuation Rosters (#7-12)		Data Defined Persons from Completed LHHFU Cases (#1-6)		All Census 2000 Persons	
	Number	Percent	Number	Percent	Number	Percent
<b>Total</b>	<b>1,327,756</b>	<b>-</b>	<b>4,363,093</b>		<b>285,230,516</b>	<b>-</b>
Male	675,500	53.0%	2,107,607	49.3%	139,887,140	49.0%
Female	600,126	47.0%	2,168,673	50.7%	145,343,376	51.0%
Missing	52,130	-	86,813	-	-	-

Source: COMBO file - HCEF variable STENURE\_HCEF

Percents are calculated excluding Missing values

**Table H3. Frequency and Percent of Age Groups of Persons on Completed Coverage Edit Followup Large household Cases Compared to the Overall Population by Persons 1-6 and Persons 7-12**

Age	Persons from LHHFU Continuation Rosters (#7-12)		Data Defined Persons from Completed LHHFU cases (#1-6)		All Census Returns	
	Number	Percent	Number	Percent	Number	Percent
<b>Total</b>	<b>1,327,756</b>	<b>-</b>	<b>4,363,093</b>	<b>-</b>	<b>285,230,516</b>	<b>-</b>
Under 5	258,246	29.4%	243,130	5.8%	19,471,204	6.8%
5 to 9 years	233,062	26.5%	515,981	12.3%	20,854,667	7.3%
10 to 14 years	118,612	13.5%	632,703	15.1%	20,833,872	7.3%
15 to 19 years	64,963	7.4%	539,315	12.8%	20,533,326	7.2%
20 to 24 years	54,310	6.2%	323,560	7.7%	19,265,192	6.8%
25 to 34 years	59,467	6.8%	567,132	13.5%	40,426,056	14.2%
35 to 44 years	30,937	3.5%	665,889	15.8%	45,664,190	16.0%
45 to 54 years	16,899	1.9%	364,887	8.7%	38,140,988	13.4%
55 to 59 years	6,970	0.8%	100,023	2.4%	13,658,120	4.8%
60 to 64 years	7,588	0.9%	77,185	1.8%	1,096,601	3.8%
65 to 74 years	14,175	1.6%	101,506	2.4%	18,631,937	6.5%
75 to 84 years	9,024	1.0%	46,545	1.1%	12,497,660	4.4%
85 years and over	3,707	0.4%	23,641	0.6%	4,287,293	1.5%
Missing or Invalid	449,796	-	161,596	-	-	-

Source: DRF2

Percents are calculated excluding Missing or Invalid values

**Table H4. Frequency and percent of race of persons on completed coverage edit followup large household cases compared to the overall population by persons 1-6 and persons 7-12**

Race	Persons from LHHFU Continuation Rosters (#7-12)		Data Defined Persons from Completed LHHFU Cases (#1-6)		All Census 2000 Persons	
	Number	Percent	Number	Percent	Number	Percent
<b>Total</b>	<b>1,327,756</b>	<b>-</b>	<b>4,363,093</b>	<b>-</b>	<b>285,230,516</b>	<b>-</b>
White	563,028	46.1%	2,250,189	71.5%	214,525,488	75.2%
Black, African American	182,800	15.0%	666,586	21.2%	34,961,123	12.3%
American Indian, Alaska Native	11,362	0.9%	50,449	1.6%	2,489,292	0.9%
Asian	77,130	6.3%	100,069	3.2%	10,250,958	3.6%
Native Hawaiian or Other Pacific Islander	8,171	0.7%	2,083	0.1%	399,928	0.1%
Some Other Race	338,268	27.7%	28,870	0.9%	15,619,084	5.5%
Two or More	39,856	3.3%	49,360	1.6%	6,984,643	2.4%
Missing	107,141	-	1,215,487	-	-	-

Source: CEFU evaluation files

Percents are calculated excluding Missing values

**Table H5. Frequency and percent of Hispanic origin of persons on completed coverage edit followup large household cases compared to the overall population by persons 1-6 and persons 7-12.**

Hispanic Origin	Persons from LHHFU Continuation Rosters (#7-12)		Data Defined Persons from Completed LHHFU Cases (#1-6)		All Census 2000 Persons	
	Number	Percent	Number	Percent	Number	Percent
<b>Total</b>	<b>1,327,756</b>	<b>-</b>	<b>4,363,093</b>		<b>285,230,516</b>	<b>-</b>
Non-Spanish/Hispanic	697,033	56.4%	2,557,794	62.0%	246,161,952	86.3%
Mexican, Mexican American, Chicano	415,627	33.6%	1,052,438	25.5%	20,652,257	7.2%
Puerto Rican	28,971	2.3%	102,459	2.5%	7,029,570	2.5%
Cuban	5,274	0.4%	18,438	0.4%	1,261,658	0.4%
Other Spanish/Hispanic	89,584	7.2%	356,299	8.6%	10,125,079	3.6%
Two or More Origins	133	0.0%	35,099	0.9%	-	-
Missing	91,134	-	240,566	-	-	-

Source: CEFU evaluation files

Percents are calculated excluding Missing values



**Appendix I: Demographics of person one for cases with ‘no valid telephone number’**

**Table I1. Frequency and percent of tenure of housing units with a final disposition of “no valid telephone number” in coverage edit followup compared to all coverage edit followup cases and the overall population**

Tenure	CEFU Cases with No Valid Telephone Number		All CEFU Cases		All Census Returns	
	Number	Percent	Number	Percent	Number	Percent
<b>Total</b>	<b>562,049</b>	<b>-</b>	<b>2,338,420</b>	<b>-</b>	<b>106,741,426</b>	<b>-</b>
Owner	305,810	55.0%	1,508,666	65.2%	70,735,522	66.3%
Renter	249,897	45.0%	805,224	34.8%	36,005,904	33.7%
Missing or Invalid	6,342	-	24,530	-	-	-

Source: DRF2

Percents are calculated excluding Missing or Invalid values

**Table I2. Frequency and percent of race of persons of person 1 on household rosters in housing units with a final disposition of “no valid telephone number” in coverage edit followup compared to all cases in coverage edit followup and the overall population**

Race	CEFU Cases with no valid telephone number		All CEFU cases		All Census Returns	
	#	%	#	%	#	%
<b>Total</b>	<b>562,049</b>	<b>-</b>	<b>2,338,420</b>	<b>-</b>	<b>106,741,426</b>	<b>-</b>
White	291,992	68.6%	1,324,503	72.8%	84,779,674	79.4%
Black, African American	106,887	25.1%	388,507	21.3%	12,159,606	11.4%
American Indian, Alaska Native	7,701	1.8%	25,478	1.4%	772,903	0.7%
Asian	10,052	2.4%	47,439	2.6%	3,132,768	2.9%
Native Hawaiian or Other Pacific Islander	201	0.0%	866	0.0%	104,281	0.1%
Some Other Race	2,170	0.5%	10,770	0.6%	3,881,418	3.6%
Two or More	6,833	1.6%	22,392	1.2%	1,910,776	1.8%
Missing or blank	136,213	-	518,465	-	-	-

Source: DRF2 and QT-H1 Summary File 1

Percents are calculated excluding Missing values

**Table I3. Frequency and percent of Hispanic origin of person 1 in housing units with a final disposition of “no valid telephone number” in coverage edit followup compared to all coverage edit followup cases and the overall population**

Hispanic Origin	CEFU Cases with No Valid Telephone Number		All CEFU cases		All Census Returns	
	Number	Percent	Number	Percent	Number	Percent
<b>Total</b>	<b>562,049</b>	<b>-</b>	<b>2,338,420</b>	<b>-</b>	<b>105,480,101</b>	<b>-</b>
Non-Spanish/Hispanic	322,195	65.9%	1,475,976	69.4%	96,257,699	91.3%
Mexican, Mexican American, Chicano	86,350	17.7%	378,806	17.8%	9,222,402*    8.7%*	
Puerto Rican	29,569	6.0%	69,540	3.3%		
Cuban	3,188	0.7%	14,613	0.7%		
Other Spanish/Hispanic	39,679	8.1%	162,738	7.6%		
Two or More Origins	7,825	1.6%	26,520	1.2%	-	-
Missing or blank	73,243	-	210,227	-	-	-

Source: DRF2 and QT-H1 Summary File 1

Percents are calculated excluding Missing values

\* Hispanic origin by householder was only available by Hispanic vs. Non Hispanic - without detailed groups

**Appendix J: Coverage questions from the 1990 and 2000 Census forms**

1990 Census coverage questions - mailback and enumerator forms

Census 2000 coverage questions - enumerator forms only

**H1a. Did you leave anyone out of your list of persons for Question 1a on page 1 because you were not sure if the person should be listed — for example, someone temporarily away on a business trip or vacation, a newborn baby still in the hospital, or a person who stays here once in a while and has no other home?**

Yes, please print the name(s) and reason(s). ↗  No

-----

**b. Did you include anyone in your list of persons for Question 1a on page 1 even though you were not sure that the person should be listed — for example, a visitor who is staying here temporarily or a person who usually lives somewhere else?**

Yes, please print the name(s) and reason(s). ↗  No

-----

**COVERAGE**

**C1. I need to make sure I have counted everyone who lived or stayed here on April 1, 2000. Did I miss —**

- any children, including foster children?
- anyone away on business or vacation?
- any roomers or housemates?
- anyone else who had no other home?

Yes - Add person(s) to question 1 and mark the "Add" box. Ask questions 2-6 for each person and correct the POP count in question S5 on the front cover.

No - Continue with C2.

**C2. The Census Bureau has already counted certain people so I don't want to count them again here. On April 1, 2000, were any of the people you told me about —**

- away at college?
- away in the Armed Forces?
- in a nursing home?
- in a correctional facility?

Yes - Delete person(s) from 1 by marking the "Cancel" box. Correct the POP count in question S5 on the front cover.

No - Continue with question 7.

ks