

# Discrepant Results in the Accuracy and Coverage Evaluation

## FINAL REPORT

This evaluation study 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.

---

Elizabeth A. Krejsa  
Planning, Research, and  
Evaluation Division

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

*Helping You Make Informed Decisions*

# CONTENTS

|  |     |
|--|-----|
| EXECUTIVE SUMMARY .....  | iii |
| 1. BACKGROUND .....  | 1   |
| 1.1 Results of previous evaluations .....                                      | 1   |
| 1.2 The A.C.E. Person Interview and Quality Assurance Reinterview .....        | 2   |
| 1.3 The Person Followup Interview and Matching Operation .....                 | 4   |
| 2. METHODS .....   | 5   |
| 2.1 Measuring Discrepant Results using the Evaluation Followup Interview ..... | 5   |
| 2.2 Questions to be Answered .....   | 6   |
| 3. LIMITS .....  | 7   |
| 4. RESULTS .....   | 8   |
| 4.1 Quality Assurance in the PI .....  | 8   |
| 4.2 Followup Studies Identifying Discrepant Results .....                      | 10  |
| 4.2.1 <i>Demographics of people identified after the QA process</i> .....      | 10  |
| 4.2.2 <i>Number of people remaining in the P-sample</i> .....                  | 10  |
| 5. RECOMMENDATIONS .....   | 12  |
| 6. REFERENCES .....  | 13  |

## LIST OF TABLES

|  |    |
|--|----|
| Table 1. Demographic Comparison of QA results and Discrepant Data .....          | 9  |
| Table 2. Impact of Change in Discrepancy Status on the Number of Residents ..... | 11 |

## **EXECUTIVE SUMMARY**

This evaluation examines how well the Accuracy and Coverage Evaluation identified discrepant results. Specifically, it studies two processes: the person interview quality assurance reinterview and the followup interview and coding operation.

Discrepant results are errors that do not include honest mistakes made by the interviewers or respondents and could be falsification but the amount is uncertain. A person is classified as discrepant during the person interview quality assurance reinterview or during the coding of the person followup interview.

The quality assurance reinterview check is conducted when:

- the housing unit is selected as part of a five percent systematic sample or
- the interviewer is targeted based on predetermined criteria and one or more of his or her interviews is selected for review.

A quality assurance interviewer returns to the housing unit and determines if the respondent, or someone else in the housing unit, was interviewed. If no one was interviewed, the quality assurance interview is designed to replace the discrepant results and the quality assurance interviewer collects data in a replacement interview for the housing unit.

The person followup interview is conducted to gain additional information on such persons, to reconcile conflicts between the Accuracy and Coverage Evaluation and Census, and to determine residency status. A person is classified as discrepant if three knowledgeable respondents indicate not knowing him or her in the person followup interview.

The evaluation followup interview, an evaluation interview collecting residence information, was conducted after the person followup interview. A person could also be identified during this interview as discrepant if three knowledgeable respondents indicated not knowing the followup person in the evaluation followup interview.

### **What were the results of the person interview quality assurance process?**

During the quality assurance process for the person interview, 637 interviewers were suspected of reporting discrepant interviews, 41 of which were confirmed by supervisors. These discrepant interviews were suspected in 979 housing units, 180 of which were confirmed by supervisors. The confirmed discrepant housing units contained 337 discrepant people who were corrected with 363 people during the quality assurance reinterview for a net increase of 26 people, or 33,811 weighted.

### **What are the differences in demographics between the discrepant and replacement persons?**

The main differences between those rostered in the discrepant interview and those rostered by the replacement interview are the addition of children and roommates as well as the addition of 18-29 year olds and those of multiple races.

**How many discrepant people were undetected by the quality assurance process as identified by the person followup and evaluation followup residence coding?**

492,017 weighted discrepant people and 65,950 weighted potentially discrepant people were identified in the followup residence coding who were not identified during the quality assurance process. Potentially discrepant people are those in which either there were not enough respondents to confirm that the person was discrepant or that the persons responding could not be sure if they knew of the person. Of the 492,017 discrepant people, 29.8 percent were in the Philadelphia Regional Office and 32.6 percent were in update/leave areas. No other demographics characteristics of the housing unit or person were different than the population distribution.

**How many discrepant people were undetected in the Accuracy and Coverage Evaluation and remain in the P-sample as identified by the evaluation followup residence coding?**

The evaluation followup identified a weighted net 326,855 residents in the P-sample who should have been removed because they were discrepant. Up to 23,879 weighted people were excluded as residents in the P-sample but identified as potentially discrepant by the evaluation followup.

## **1. BACKGROUND**

This evaluation examines how well the Accuracy and Coverage Evaluation (A.C.E.) identified discrepant results. Specifically, it studies two processes: the person interview quality assurance reinterview and the matching operation.

Discrepant results are errors that do not include honest mistakes made by the interviewers or respondents and could be falsification but the amount is uncertain. A person is classified as discrepant during the production matching operation if three knowledgeable respondents indicate not knowing him or her in the Person Followup interview (PFU) (Martinez, 2001).

In missing data processing the discrepant people are assigned zero probability of correct enumeration in the enumeration sample (E-sample) or are assigned zero probability of being a resident in the population sample (P-sample) (Childers, 2000). If discrepant people were not identified as discrepant in the A.C.E. then the estimate of correct enumerations is inflated in the E-sample as well as in the estimate of the resident population in the P-sample.

### **1.1 Results of previous evaluations**

In past censuses and tests, similar evaluations were conducted.

Three evaluations of the followup survey to the 1990 Census were conducted on discrepant results. This followup survey known as the Post-Enumeration Survey (PES) was an independent enumeration similar to the A.C.E. West examined discrepant results in the P-sample using a questionnaire to gather information to resolve residence status and match code conflicts between PES and Census data. The data were collected in February 1991 in approximately 11,000 housing units nationwide to determine residence status on Census Day. Thirteen people, not previously identified as discrepant, were coded by matchers as discrepant in the evaluation followup matching operation. The thirteen people represented 64,667 people, or .03 percent of the total estimated population. The standard error of that estimate is 39,419 (West, 1991a).

West also examined discrepant results in the P-sample by exploring an interviewer effect on the data. To examine the effectiveness of the quality control (QC) procedures a model was developed to predict the nonmatch rate obtained by interviewers. Standardized scores (Z-scores) were computed for each interviewer and deviations from the model were used to suggest discrepancies. The results showed that a number of interviewers had different nonmatch rates from other interviewers working in similar areas. Between 0.7 percent and 5.4 percent of the total completed workload was not identified during the QC operation (West, 1991b).

Tremblay analyzed P-sample discrepant results from the QC data. The report provides descriptive statistics of those households for which the QC check detected the entire household

roster to be in error (that is, discrepant households that were caught and replaced). It also determines residual discrepancies (that is, discrepant results remaining in the P-sample after the QC operation). Tremblay estimated that 336 weighted households contained discrepant persons. Weighted to the national level, she estimated 310,050 weighted persons remained in the P-sample after the QC operation, approximately .13 percent of the population (Tremblay, 1991).

In the Census 2000 Dress Rehearsal it was found that between 0.0 and .06 percent of the total person interviews were discrepant in the three dress rehearsal sites. It was also determined that the procedures used to identify discrepant people were satisfactory; this procedure, with some improvements (such as fixes in the questionnaire), was used for Census 2000 (Krejsa, 1999). No estimate of the residual discrepancies was calculated.

## **1.2 The A.C.E. Person Interview and Quality Assurance Reinterview**

As part of the A.C.E. operation, a person interview (PI) is conducted. To help ensure that the data collected are valid, a quality assurance (QA) reinterview is done. The purpose of the QA reinterview is to determine if discrepant results were entered by the interviewer during the original A.C.E. PI. The QA reinterview check is conducted when:

- the housing unit is selected as part of a five percent systematic sample or
- the interviewer is targeted based on predetermined criteria and one or more of his or her interviews is selected for review.

A QA interviewer returns to the housing unit and determines if the respondent, or someone else in the housing unit, was interviewed. If no one was interviewed, the QA is designed to replace the discrepant results and the QA interviewer collects a replacement interview for the housing unit. A replacement interview is conducted when the PI interviewer is suspected of reporting discrepant results. The QA supervisor in the A.C.E. regional office makes the final decision whether an interviewer has failed QA. To determine if the interviewer has failed QA by reporting discrepant results, the QA supervisor will speak to the QA interviewer, the PI respondent, and possibly the PI interviewer (rarely).

### How a Housing Unit is Systematically Selected for QA

A five percent systematic sample of housing units is selected before any interviewer assignments are made for the PI. Any PI noted as 'a language problem', 'no respondent to conduct the interview', or 'respondent refusal' does not go to QA, even if selected as part of the systematic sample.

### How an Interviewer is Targeted for QA

The goal in targeting interviewers is to make the QA effort more efficient and effective in identifying discrepant results and interviews of poor quality (such as proxy or partial interviews).

Targeting refers to the identification of interviewers who are possibly recording discrepant results or conducting poor quality interviews and sending part of these interviewers' workload to QA reinterview. Targeting reports are used to assist the QA staff in selecting interviewers' cases. The reports contain statistics believed to be good indicators of problem interviews and are intended to assist QA supervisors in targeting suspicious interviewers and interviews for QA reinterview. The systematically selected housing units are not eligible to be targeted. There are four targeting reports.

- Respondent's Name Report

This report lists the entries for the respondent names by each interviewer. If an interviewer is reporting discrepant results, he or she may have difficulty thinking of respondent names. Sometimes he or she will use names of famous people or cartoon characters, or repeat a name in more than one interview. It was also noted that interviewers were using the names used as examples during training. The QA supervisors look for such situations.

- Field Representative Outlier Report

This report identifies interviewers who have outliers, compared to other interviewers in the same work area, for characteristics related to discrepant results and data quality. Tolerance levels are predetermined for some characteristics. Any case that falls within one of these levels is indicated and can be sent for a QA reinterview. The outlier characteristics with predetermined tolerance levels are:

- cases completed after regular hours (between 10 p.m. and 8 a.m.)
- more than thirteen completed interviews in a day

For other characteristics the tolerance levels are based on the average for the work area. An interviewer identified as an outlier for these characteristics is flagged and the QA supervisor selects a case with the given characteristic for QA reinterview, unless the QA supervisor can explain why the particular interviewer would be an outlier. The outlier characteristics are:

- a. percent of cases with no phone number
- b. percent of cases with missing outmover data where a noninterview occurred for Census Day but an interview was completed or a vacant housing unit was found on interview day for the PI
- c. percent of vacant housing units and housing units that do not exist or are no longer used for residential purposes (i.e., now is a business) on interview day, not Census day
- d. percent of cases completed by a proxy respondent
- e. percent of cases in which some data are missing for at least one person in the housing unit

Characteristics b, d, and e are used as quality indicators to help identify interviewers who may be following incorrect procedures. All other characteristics are used as discrepancy indicators.



Tolerance levels based on averages are calculated for a specific work area to determine outliers. It is assumed that cases within a work area are of similar difficulty for the interviewers and have similar characteristics. The tolerance level used to determine the outliers is calculated by the following:

$$\text{Outlier: } p_{FR} > p_{pop} + 1.645 * (\sigma / \sqrt{n})$$

where: FR is the interviewer (they are synonymous terms)

$p_{FR}$  is the interviewer proportion of cases with the characteristic

$p_{pop}$  is the proportion of cases with the characteristic currently completed and sent back from field within a work area (population)

$n$  is the number of interviewers in the given work area

$\sigma = \sqrt{p_{pop}(1-p_{pop})}$  is the population standard deviation

- Not Enough QA Cases Report

This report enables the QA supervisor to identify interviewers who have completed and sent back at least ten interviews but have no cases in QA. The QA supervisor then randomly selects one case and sends it to QA. This report is especially helpful at the beginning of the QA process when there are not a lot of cases to generate the outlier reports.

- Add QA Cases

This report is used if a supervisor identifies a potential problem with an interviewer's work in one of the other reports and wants to assign specific cases or additional cases to QA. In addition, once an interviewer is identified as having possibly reported discrepant results for part of his or her workload, additional cases are added to the QA reinterview for verification. If it is confirmed that the interviewer reported discrepant results, the interviewer has failed QA and as many interviews as possible conducted by that interviewer are redone. For further details on the QA plan see Byrne, 1999a, Byrne, 1999b, and Byrne, 1999c.

### **1.3 The Person Followup Interview and Matching Operation**

If the QA reinterview fails to detect discrepant results, they may be identified from the Person Followup (PFU) interview and the matching operation. The Census Bureau attempts to match people collected in the Census to people collected in the PI. Sometimes, additional information is needed to match. The PFU interview is conducted to gain additional information on such persons, to reconcile conflicts between A.C.E. and Census, and to determine residency status. Housing units and persons sent to PFU include:

- A.C.E. non-matched persons in housing units in which some of the A.C.E. people match to Census people and some do not,

- A.C.E. non-matched persons in housing units where none of the people match and the PI was conducted with a proxy,
- A.C.E. non-matched persons in housing units where none of the people match to the persons in the Census, where none of the people in the Census housing unit match to anyone in the A.C.E., but the housing unit in the A.C.E. listing matched the housing unit in the Census listing,
- Census people who did not match to A.C.E. people,
- A.C.E. and Census persons who may be matched but further information is needed, and
- A.C.E. people whose residency status for April 1<sup>st</sup> cannot be determined.

After the PFU is conducted, another step in the matching operation takes place; this matching process is called After Followup (AFU). It is during AFU matching that people may be classified as discrepant. Therefore, these people were not identified in the QA process. A person is classified as discrepant if three knowledgeable respondents indicate not knowing him or her in the PFU interview. The three knowledgeable respondents must have answered “No” to the questions,

- Do you know or have you heard of (the A.C.E. person)?
- Do you know someone else who might know (the A.C.E. person)?

Once a person is classified as discrepant he or she is flagged on the file. In missing data processing these people are assigned zero probability of residence and match (Childers, 2000). In estimation, because of their zero probabilities, they drop out of the final estimates.

## 2. METHODS

### 2.1 Measuring Discrepant Results using the Evaluation Followup Interview

To evaluate the identification of discrepant results in the A.C.E., a personal visit dependent reinterview, called the Evaluation Followup Interview (EFU), was conducted. The EFU is conducted in approximately 1-in-5 A.C.E. clusters. These clusters are called the evaluation cluster sample. The evaluation clusters are selected by first stratifying the A.C.E. clusters by region (Northeast, Midwest, South, West), minority status (Hispanic minority, Non-Hispanic minority, Non-minority), problem status (problem, nonproblem), and cluster size (these are sorted, not stratified). Problem status is assigned to clusters having the largest five percent of a “score”. The “score” is calculated to reflect the weighted number of P-sample nonmatches, the weighted number of P-sample and E-sample unresolved persons, and the weighted number of erroneous enumerations (Childers, 2001). Sampling rates differ by stratum. All clusters within a stratum have equal probabilities of selection. Clusters are selected with certainty in the problem strata. A 1-in-4 sampling rate is used in the minority-by-non-problem strata. Clusters in the non-minority-by- non-problem strata are selected using an appropriate sampling rate; this sampling rate (greater than 1-in-5) results in meeting the goal of selecting an overall sample of 1-in-5 A.C.E. clusters (Keathley, 2001).

The EFU data were collected from mid-January to mid-February, 2001 encompassing 31,513 housing units in 2,259 evaluation clusters. Most housing units that were sent to the production PFU interview are included in the EFU field work. In addition the following persons and housing units were followed up in EFU:

- a sample of A.C.E. housing units in which all persons match to the Census,
- a sample of A.C.E. housing units where none of the people match to the Census and the PI was conducted with a housing unit member,
- all matched people in housing units where only some of the people matched, and
- housing units in which someone was identified as having moved in between April 1, 2000 and the PI interview and there were census day residents. Housing units with in-movers where there were no census day residents were inadvertently excluded from the EFU.

The purpose of the EFU is similar to the PFU in that it gathers information to resolve conflicts between A.C.E. and Census and to determine residence status. In addition, the EFU identifies reasons why a person may have been erroneously listed or not listed as a census day resident in the A.C.E. or Census. This is done by collecting the dates a person moved in or out of the housing unit during the year 2000, determining if the person is a college student and where he or she lived while attending college, determining if the person was in a group quarters on census day, and determining if the person was actually living at another residence on census day, such as one used while working away from home.

After the EFU interview was completed, a clerical matching operation took place. The EFU matching operation is similar to the AFU matching operation except the matchers use the information from the EFU interview to determine true residence status on census day. A person was classified as discrepant in this operation if three knowledgeable respondents indicated not knowing him or her in the EFU interview, in the same way a person was classified as discrepant in PFU. For further details on the EFU interview and matching operation see “Study Plan for O.14: Analysis of Measurement Error” by David Raglin, February 9, 2001.

## **2.2 Questions to be Answered**

This evaluation examines discrepant results by determining how well the A.C.E. PI quality assurance process did in identifying discrepant results by answering the following questions:

- a. How many interviewers were suspected of reporting discrepant results as identified by the PI QA? How many housing units were suspected of containing discrepant people as identified by the PI QA?
- b. How many interviewers were confirmed to have failed QA as identified by the QA supervisors? How many housing units does this include as identified by the QA supervisors? What are the demographics of the discrepant persons?
- c. How many discrepant people were undetected by the QA process as identified by the AFU and EFU matching?

- d. How many discrepant people were undetected in the A.C.E. and remain in the P-sample as identified by the EFU matching?

### **3. LIMITS**

Studies of discrepant results for surveys in general estimate that between 2 and 5 percent of interviewers may report discrepant results and that between .5 and 1.5 percent of the interviews may be discrepant (Stokes and Jones, 1989). Since it is rare for interviewers to provide discrepant results, the QA operation, the AFU matching operation, and the EFU matching operation will provide few cases. Therefore, conclusions and/or recommendations from this evaluation are drawn with caution. In addition, due to the small amount of data some originally planned tabulations are not computed and hypothesis tests not carried out.

While QA procedures for the PI were set up in advance of the operation, several operational quirks occurred that obscure the actual number of housing units believed to contain discrepant persons and interviewers who reported discrepant results. For example, some regions reassigned the completed PI cases for persons who are believed to have reported discrepant results rather than adding their cases to QA. Thus QA results for those reassigned cases are unavailable because they did not go to QA. The supervisors' reports and notes from the interviews were heavily utilized to determine a more accurate estimate.

The code used in the AFU matching operations identifies discrepant persons. These persons may be discrepant due to interviewer falsification or other types of interviewer error or due to respondent error. Therefore, the matching code is not equivalent to the confirmation of an interviewer failing QA.

Because the EFU interview takes place 9 months after Census Day there is concern that information reported in the EFU may not be as valid as the information received during the production process. For this reason, matchers are given the option to reject the information received in the EFU interview in favor of the production match code. Information obtained for roughly 10 percent of people who were followed up in EFU was rejected. In general, if information is obtained about a person in the PFU interview but not in the EFU interview, the EFU interview for that person is rejected and the production match code is kept, as opposed to coding the person discrepant in the EFU (Green, 2001).

Another limitation is that the EFU did not have a full field quality assurance program as did the A.C.E. Person Interview and the PFU.

## 4. RESULTS

### 4.1 Quality Assurance in the PI

The number of interviewers suspected of reporting discrepant results and number housing units suspected of containing discrepant people is calculated based on the results of the QA interview. Suspected discrepancies are identified by combining two indicators: if the respondent says that he or she (or someone else in the housing unit) had not been interviewed and if the QA interviewer believes the original interviewer reported discrepant results based on the results of the interview. Based on this assessment, 5.7 percent of interviewers (637 interviewers of the 11,132 who worked the PI operation) were suspected of reporting discrepant results. Roughly 52.2 percent of those interviewers were suspected of reporting discrepant results for only one interview. The PI QA process identified .3 percent (979 housing units of the 300,913 housing units in the A.C.E.) of housing units suspected of containing discrepant people.

Any case suspected of containing discrepant results is reviewed by supervisors to determine if the original interviewer actually reported discrepant results. This information is recorded on a supervisor's report and is examined to calculate how often interviewers failed QA. Forty-one interviewers were confirmed to have reported discrepant data. These interviewers worked in ten of the twelve regions. This calculation is highly subjective however. Some regions never reported an interviewer as confirmed; others were less conservative with their judgements. Discrepant people were confirmed in 180 housing units of the 979 housing units suspected of containing discrepant people. These discrepant housing units were reported to contain 337 people. However, when the interviews were replaced, a total of 363 people were rostered yielding a net increase of 26 people. To determine the weighted net increase we first calculated the weighted number of people in the QA interview using the assigned weights. We then calculated what the weighted number of people would have been if the discrepant interview had not been detected. These people were not assigned weights so the average cluster weight was used. Weighted to the national level, these 26 people represent an increase of 9,908 people. Below is a summary, at the national level, of how the demographics changed when the confirmed discrepant people were replaced.

Table 1. Demographic Comparison of QA results and Discrepant Data

| Demographic Description and Category |                           | Number Reported in the QA Interview | Number Reported in the Discrepant Interview |
|--------------------------------------|---------------------------|-------------------------------------|---|
| Sex                                  | Male                      | 170                                 | 152   |
|                                      | Female                    | 168                                 | 158   |
| Age                                  | 0-17                      | 83                                  | 49  |
|                                      | 18-29                     | 105                                 | 77  |
|                                      | 30-49                     | 95                                  | 107   |
|                                      | 50+                       | 47                                  | 44  |
| Race                                 | White                     | 177                                 | 167   |
|                                      | Black                     | 78                                  | 74  |
|                                      | Native American           | 5                                   | 7   |
|                                      | Asian/Pacific Islander    | 14                                  | 11  |
|                                      | Other                     | 43                                  | 42  |
|                                      | Multiple Races            | 12                                  | 0   |
| Hispanic Origin                      | Yes                       | 60                                  | 64  |
|                                      | No                        | 279                                 | 250   |
| Relationship                         | Husband/Wife              | 51                                  | 56  |
|                                      | Child                     | 91                                  | 44  |
|                                      | Other type of Relative    | 12                                  | 9   |
|                                      | Roommate                  | 20                                  | 3   |
|                                      | Other type of Nonrelative | 8                                   | 11  |
| Housing unit Size                    |                           | 2.02                                | 1.87  |
| Tenure                               | Occupied, Owned           | 48                                  | 59  |
|                                      | Occupied, Rented          | 74                                  | 98  |
|                                      | Occupied without payment  | 1                                   | 2   |
|                                      | Occupied, tenure unknown  | 54                                  | 3   |
|                                      | Vacant Unit               | 0                                   | 15  |

Answers of ‘don’t know’ and ‘refuse’ were rarely given for either interview and do not appear in the table above.

The most notable change in demographics occurred in the rostering of children and roommates in housing units in the QA interview versus the interview containing discrepant results. The additional rostering of children and roommates is logical. When an interviewer reports discrepant data they may use information from a telephone book or mailbox. Such information is likely to exclude children or roommates (since typically only one name is listed in a phone book) and thus these people would be omitted from the roster. Other notable differences were an increase in the number of 18-29 year olds reported and multiple races. Also, fifteen units were reported vacant in the discrepant interview; there were no vacant units reported by the QA interviewer for any of the units.

## 4.2 Followup Studies Identifying Discrepant Results

Discrepant results detected after the QA process, through the AFU matching and through the EFU matching, are used to calculate residual discrepancies in the P-sample. To determine how many discrepant people were undetected by the QA process, a combination of the AFU and EFU results for those cases in the 2,259 evaluation clusters is used. The non-interview adjustment makes some weights zero as a result of identifying discrepant persons. To avoid the concealment of identification of such discrepancies the weight used to calculate the figures below excludes the non-interview adjustment.

### 4.2.1 *Demographics of people identified after the QA process*

The AFU and EFU matching processes identified 492,071 people (257 unweighted) as being discrepant persons with an additional 65,950 people (63 unweighted) identified as potentially discrepant. Potentially discrepant people are those in which either there were not enough respondents to confirm that the person was discrepant or that the persons responding could not be sure if they knew of the person. In addition, the coding of discrepant persons in the matching operations was difficult. If in AFU matching someone was coded discrepant but in the evaluation matching was coded unresolved the person was considered potentially discrepant.

Workloads for the regional offices were roughly evenly distributed (ranging from 6.35 percent of the workload to 11.21 percent). However, the distribution of missed discrepant cases was not evenly distributed across the regional offices. Of the 492,071 people identified after the QA process, 29.8 percent were in the Philadelphia Regional Office (which had 7.14 percent of the national workload), 14.3 percent were in the Charlotte Regional Office (which had 11.21 percent of the national workload), and 11.9 percent were in the Seattle Regional Office (which had 7.77 percent of the national workload). In addition, 32.6 percent were in update/leave areas (which was conducted in 17.37 percent of the country).

Most demographic characteristics of these people represented the distribution of the people in the same clusters. There was a slightly higher rate of 18-29 year olds, males, and non-relatives identified as discrepant after the QA process than found in the population.

#### 4.2.2 Number of people remaining in the P-sample

To determine how many people were undetected in the A.C.E. and remain in the P-sample, those people who were newly identified as discrepant persons by the EFU matching operation are totaled. Of the 492,071 persons identified after QA as being discrepant, 98,196 (85 unweighted) were identified in the production process. Therefore, 393,876 people (172 unweighted) identified by the EFU were included in the P-sample who should have been identified as discrepant.

While the count of residual discrepant persons gives us insight into differences in matching and followup information obtained, this number does not adequately represent the impact on the number of residents for the dual system estimator. When someone is determined to be discrepant, they have a zero probability of being a resident in the cluster on Census Day. When someone is potentially discrepant there is a probability of being a resident greater than zero but typically lower than one-half. Table 2 below shows the number of residents determined by production and the evaluation.

Table 2. Impact of Change in Discrepancy Status on the Number of Residents

|  | Weighted* Total      | Evaluation Residents | Production Residents |
|--|----------------------|----------------------|----------------------|
| Discrepant in Production and Evaluation  | 98,196<br>(21,640)   | 0                    | 0                    |
| Discrepant in Evaluation only            | 393,876<br>(133,175) | 0                    | 360,122<br>(132,194) |
| Discrepant in Production, not Evaluation | 58,249<br>(12,186)   | 33,267<br>(8,578)    | 0                    |
| Potentially Discrepant                   | 65,950<br>(16,386)   | 30,968<br>(8,070)    | 7,089<br>(2,740)     |

\* The non-interview adjustment weight is excluded in this table.

Someone who was determined to be discrepant according to production would not be counted as a resident. There were 58,249 people that production considered discrepant but the evaluation was able to determine existed in the cluster. Of these 58,249 people, 33,267 were determined to be census day residents in the evaluation; therefore the production process erroneously excluded these people. However, there were 393,876 people who production coded as being non-discrepant people but who the evaluation determined were discrepant. Of these 393,876 people, 360,122 were included but should not have been.

The potentially discrepant cases are those cases which may be discrepant but we cannot determine. In these types of cases a judgement was made by the coder in each operation about whether the three respondents were considered knowledgeable enough about the housing unit to know if someone was discrepant or not. If the coder did not believe the respondent was



knowledgeable an unresolved code was applied. If the coder did, then a discrepant code was applied. Since this judgement could vary among the coders, it seems reasonable that either the evaluation or the production code could be true. This was the case for 65,950 people. Because it is difficult to determine if the respondent is in a position to be knowledgeable about the person in question and to account for the coding differences, unresolved people are imputed for at a very low rate of residence. If we believe the data from the evaluation 30,968 people should be included in the residence count; if we believe the production data 7,089 people should be included.

These figures, however, are larger than the true impact on the residence count because the non-interview adjustment is excluded.

## **5. RECOMMENDATIONS**

The PI Quality Assurance process successfully identified the reporting of discrepant persons during the PI operation. However, additional discrepant persons were identified in post PI operations. There was no specific commonality among these interviewers that would have identified them earlier as providing discrepant results. The practice of conducting a QA check on each interviewer at least once during the PI operation appears to be an effective method of detecting discrepant interviews and should continue to be used.

## 6. REFERENCES

Byrne, Rosemary (1999a). "Census 2000 Quality Assurance Plan for the Accuracy and Coverage Evaluation CAPI Person Interview." DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-QA-03. May 17, 2000.

— (1999b). "Census 2000 Accuracy and Coverage Evaluation QA Targeting Reports Specifications for Person Interviews." DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-QA-06. Working Draft. October 23, 2000.

— (1999c). "Census 2000 Specifications for Training of QA Supervisors on QA Targeting Reports." DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-QA-07. October 23, 2000.

— (1999d). "Specifications for the Quality Assurance CAPI Instrument for Person Interviewing in the A.C.E; Revision 2." DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-IT-01. June 21, 1999.

Coverage Measurement Staff. "Evaluations Operational Plan for the Census 2000 Accuracy and Coverage Evaluation." PRED TXE/2010 Memorandum Series: CM-GES-S-03. January 15, 2003.

Childers, Danny R (2000). "Accuracy and Coverage Evaluation: The Design Document." DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-DT-01. October 11, 2000.

— (2001). "Algorithm for Determining Outlier Clusters in the 2000 A.C.E." DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-DT-06. February 31, 2001.

Green, Alicia. "Measurement Error Reinterview Technician Manual (D-1500M)."

Keathley, Don H. "EFU Sample Design, Stratification, Selection, and Weighting." PRED TXE/2010 Memorandum Series: CM-GES-S-02. February 26, 2001.

Krejsa, Elizabeth A. "Evaluation of the Quality Assurance Falsification Model for the Integrated Coverage Measurement Person Interview." Census 2000 Dress Rehearsal Evaluation Results Memorandum Series #C5. U.S. Bureau of the Census. July 1999.

Martinez, Marjorie. "After Follow-Up Person Clerical Manual (D-1825)." Census 2000 Procedures and Operations Memorandum Series, Chapter S-TM-11. June 15, 2001.

Raglin, David. "Study Plan for O.14: Analysis of Measurement Error." Census 2000 Evaluation Program. February 9, 2001.

Stokes, S.L. and Jones, P.M. (1989) "Evaluation of the Interviewer Quality Control Procedure for the Post-Enumeration Survey," American Statistical Association. 1989 Proceedings of the Section on Survey Research Methods, 696-698.

Tremblay, Antoinnette. "Evaluation Project P5: Analysis of PES P-sample Fabrications from PES Quality Control Data". 1990 Coverage Studies and Evaluation Memorandum Series #E-4. U.S. Bureau of the Census. July 9, 1991.

West, Kirsten K (1991a). "1990 Post-Enumeration Survey Evaluation Project P5a: Analysis of P-sample Fabrication from Evaluation Followup Data". STSD Coverage Studies and Evaluation Memorandum Series #F-1. U.S. Bureau of the Census. July 9, 1991.

---(1991b). "1990 Post-Enumeration Survey Evaluation Project P6: Fabrication in the P-sample: Interviewer Effect". 1990 Coverage Studies and Evaluation Memorandum Series #G-2. U.S. Bureau of the Census. July 10, 1991.