United States General Accounting Office

GAO

Report to Congressional Requesters

March 2002

2000 CENSUS

Coverage Evaluation Matching Implemented as Planned, but Census Bureau Should Evaluate Lessons Learned

G A O

Accountability ★ Integrity ★ Reliability

GAO-02-297

Contents

Letter

1

Results in Brief 2
Background 4
Matching Process Was Complex, and Application of Criteria Involved the Judgment of Trained Bureau Staff 6
Quality Assurance Results Suggest Person Matching Procedures Were Implemented as Planned 12
The Bureau Took Action to Address Some Deviations, but Effect on Matching Results Is Unknown 16
Conclusions 20
Recommendations for Executive Action 21
Agency Comments and Our Evaluation 21

Appendixes

Appendix I: Scope and Methodology 24
Appendix II: Comments from the Department of Commerce 25
Appendix III: GAO Contact and Staff Acknowledgments 28

Table

Table 1: Deviations from the Planned Person Matching Operation 17

Figures

Figure 1: A.C.E. Survey Followed Steps Similar to Census 5
Figure 2: Person Matching, Quality Assurance Coverage 14
Figure 3: Quality Assurance of Field Follow-up by A.C.E. Regional Office 15

G A O
Accountability * Integrity * Reliability

United States General Accounting Office
Washington, D.C. 20548

March 14, 2002

The Honorable Dave Weldon
Chairman
The Honorable Danny K. Davis
Ranking Minority Member
Subcommittee on Civil Service, Census and Agency Organization
Committee on Government Reform
House of Representatives

The Honorable William Lacy Clay
The Honorable Carolyn B. Maloney
The Honorable Dan Miller
House of Representatives

To assess the quality of the population data collected in the 2000 Census, the U.S. Census Bureau conducted the Accuracy and Coverage Evaluation (A.C.E.) survey, a sample of persons designed to estimate the number of people missed, counted more than once, or otherwise improperly counted in the census. On the basis of uncertainty in the A.C.E. results, in separate decisions in March and October 2001, the acting director of the bureau decided that the 2000 Census tabulations should not be adjusted for purposes of redrawing the boundaries of congressional districts or for other purposes, such as distributing billions of dollars in federal funding. Although A.C.E. was generally implemented as planned, the bureau found that A.C.E. overstated census undercounts due in part to error introduced during matching operations and other remaining uncertainties. The bureau has reported that additional review and analysis on these remaining uncertainties would be necessary before any potential uses of these data can be considered.

A critical component of the A.C.E. survey was the person matching operation, in which the bureau matched the persons counted in the A.C.E. survey to the persons counted in the census. The results of person matching formed the basis for statistical estimates of the proportions of the population missed or improperly counted by the census.

This report, prepared at the request of the chairman and ranking minority member of the former House Subcommittee on the Census, reviews the person matching operation of A.C.E. We agreed to describe (1) the process and criteria involved in making an A.C.E. and census person match, (2) the quality assurance procedures used in the key person matching phases and

the available results of those procedures, and (3) any deviations in the matching operation from what was planned. This report is the latest of several we have issued on lessons learned from the 2000 Census that can help inform the bureau's planning efforts for the 2010 Census.

To address our three objectives, we examined relevant bureau program specifications, training manuals, office manuals, memorandums, and other progress and research documents. We also interviewed bureau officials at bureau headquarters in Suitland, Md., and the bureau's National Processing Center in Jeffersonville, Ind., which was responsible for the planning and implementation of the person matching operation. Further scope and methodological details are given in appendix I. We performed our audit work from September 2000 through April 2001 in accordance with generally accepted government auditing standards. On January 4, 2002, we requested comments on a draft of this report from the secretary of commerce. On February 13, 2002, the secretary of commerce forwarded written comments from the bureau (see appendix II), which we address in the "Agency Comments and Our Evaluation" section of this report.

Results in Brief

Matching over 1.4 million census and A.C.E. records was a complex and often labor-intensive process that consisted of four phases, each with its own matching procedures and multiple layers of review. The four phases were as follows.

- Computer matching, which took pairs of A.C.E. and census records and compared certain personal characteristics such as last name and age. The computer assigned a match score to each pair of records based on the extent to which the characteristics aligned. Experienced bureau staff then judgmentally determined cutoff scores to separate the groups of records that would be coded as a "match," "possible match," or one of a number of codes that defines them as not matched. However, bureau staff did not document the criteria they used to determine the cutoffs. As a result, future bureau staff may not benefit from the lessons learned by current staff about how cutoff scores are applied.
- Clerical matching (first phase), in which over 250 trained bureau staff reviewed all records and attempted to link those records left unmatched in the previous phase, in part by matching records that contained abbreviations and spelling differences.
- Field follow-up, in which bureau interviewers visited households where additional information was needed to assign match codes to a pair of records.

- Clerical matching (second phase), in which clerks used information obtained from field follow-up to match and conduct a final review of records. The bureau coded as "unresolved" records without enough information to be coded otherwise. The bureau then used statistical imputation methods to assign a match code to records coded as "unresolved," based on an examination of the results of similar records for which the bureau was able to assign a match code. While some imputation is unavoidable, it introduces uncertainty into the estimates of census over- or undercount rates.

The bureau applied quality assurance procedures to each phase of person matching. For example, during the field follow-up phase, supervisors and office staff were to review each questionnaire for legibility and completeness. In addition, A.C.E. regional offices were to reinterview a random sample of 5 percent of the households to ensure that enumerators had not falsified data. Because the quality assurance procedures had failure rates of less than 1 percent, the bureau reported that person matching quality assurance was successful at minimizing errors.

Overall, the bureau carried out person matching as planned, with few procedural deviations. The operation deviated somewhat from what was planned as a result of programming errors, printing problems, and events that triggered delays. Although the bureau addressed these deviations and person matching continued, in some cases the effect the deviations had on person matching is unknown. For example, because of printing and other problems, pages and names were missing from some of the follow-up questionnaires, and a section that verified whether the person being matched was in the geographic sample area was incomplete in some others. The bureau was unable to document the extent, effect, or cause of the printing problems and coded incomplete questionnaires as "unresolved." Bureau officials believe that the effect of the deviations was small based on the timely actions taken to address them. Nevertheless, although the bureau has concluded that A.C.E. matching quality improved compared to that in 1990, the bureau has reported that matching error remained and contributed to an overstatement of the A.C.E. estimate of census undercounts. Furthermore, despite the improvement in matching reported by the bureau, A.C.E. results were not used to adjust the census because of these errors as well as other remaining uncertainties. Therefore, it will be important for the bureau to determine the impact of these operational deviations.

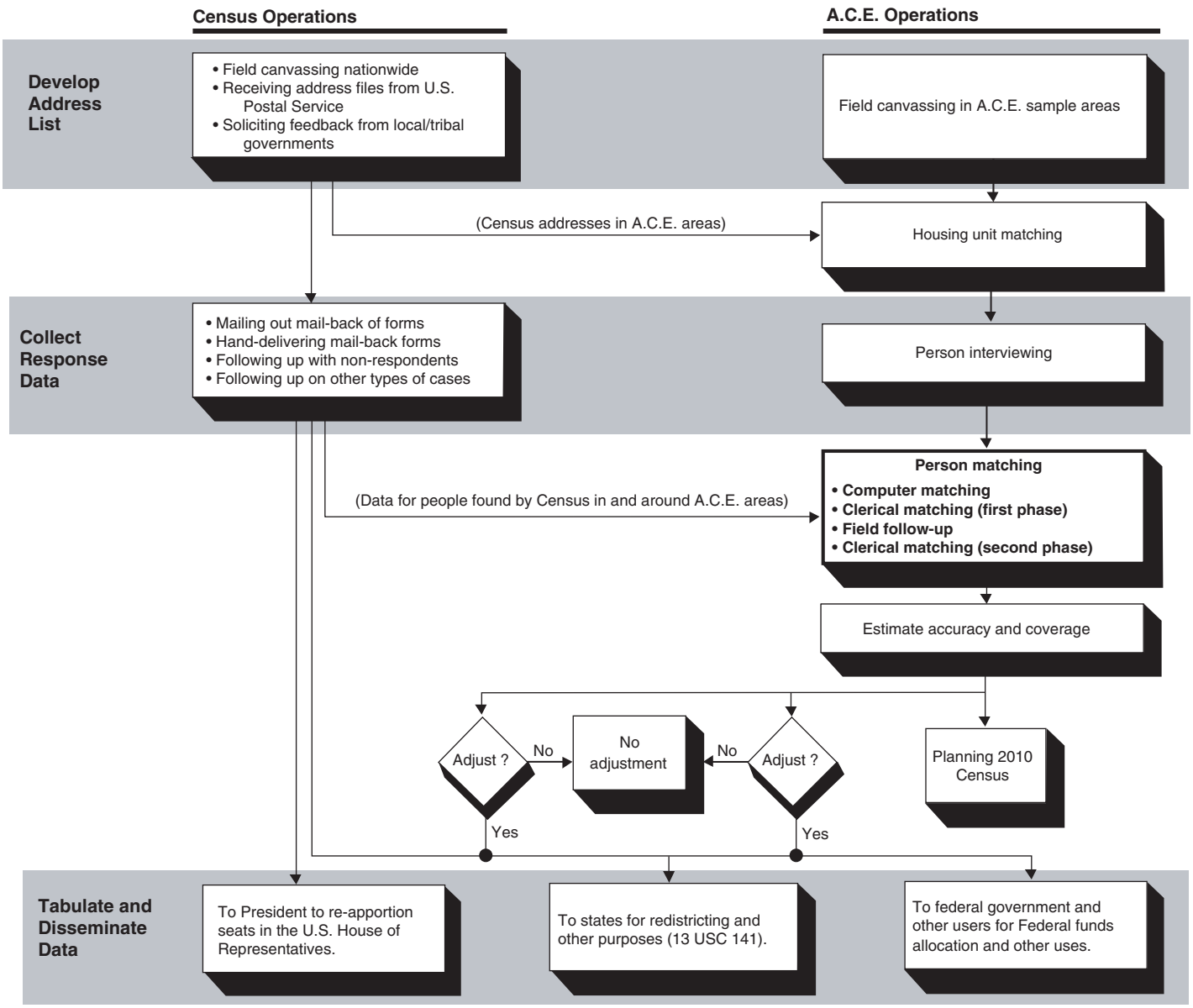
Our review identified areas with opportunity for improving future A.C.E. efforts, including more complete documentation of computer matching decisions and better assurance that problems do not arise with the bureau's automated systems. Therefore, as part of the bureau's effort to isolate lessons learned from the 2000 Census and to prepare for the census in 2010, we recommend that the secretary of commerce direct the bureau to (1) document the criteria used during computer matching to determine the groups of matched, possibly matched, and nonmatched records, (2) determine why problems with some of its automated systems were not discovered prior to deployment, and (3) determine the effect that deviations from planned operations may have had on the matching results for affected records and thus the accuracy of A.C.E. estimates of census undercounts.

The secretary of commerce forwarded written comments from the U.S. Census Bureau on a draft of this report. (See appendix II.) The bureau had no comments on the text of the report and agreed with, and is taking action on, two of our four recommendations. The bureau provided additional clarification on our other two recommendations. We comment further on the bureau's response in the "Agency Comments and Our Evaluation" section of this report.

Background

From April 24 through September 11, 2000, the U.S. Census Bureau surveyed a sample of about 314,000 housing units (about 1.4 million census and A.C.E. records in various areas of the country, including Puerto Rico) to estimate the number of people and housing units missed or counted more than once in the census and to evaluate the final census counts. Temporary bureau staff conducted the surveys by telephone and in-person visits. The A.C.E. sample consisted of about 12,000 "clusters" or geographic areas that each contained about 20 to 30 housing units. The bureau selected sample clusters to be representative of the nation as a whole, relying on variables such as state, race and ethnicity, owner or renter, as well as the size of each cluster and whether the cluster was on an American Indian reservation. The bureau canvassed the A.C.E. sample area, developed an address list, and collected response data for persons living in the sample area on Census Day (April 1, 2000). Although the bureau's A.C.E. data and address list were collected and maintained separately from the bureau's census work, A.C.E. processes were similar to those of the census.

Figure 1: A.C.E. Survey Followed Steps Similar to Census



Source: U.S. Census Bureau documents.

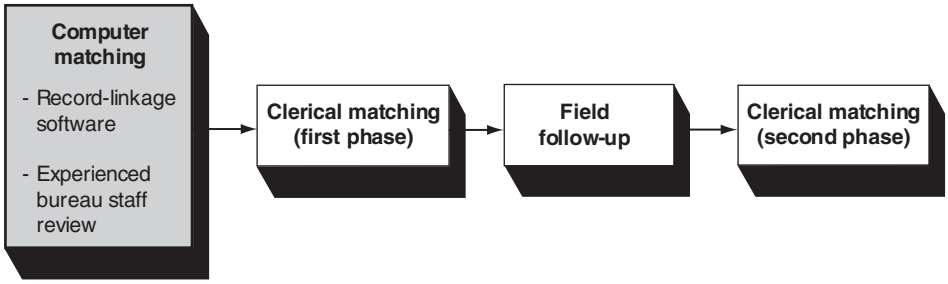
After the census and A.C.E. data collection operations were completed, the bureau attempted to match each person counted by A.C.E. to the list of

persons counted by the census in the sample areas to determine the number of persons who lived in the sample area on Census Day. The results of the matching process, together with the characteristics of each person compared, provided the basis for statistical estimates of the number and characteristics of the population missed or improperly counted by the census. Correctly matching A.C.E. persons with census persons is important because errors in even a small percentage of records can significantly affect the undercount or overcount estimate.

Matching Process Was Complex, and Application of Criteria Involved the Judgment of Trained Bureau Staff

Matching over 1.4 million census and A.C.E. records was a complex and often labor-intensive process. Although several key matching tasks were automated and used prespecified decision rules, other tasks were carried out by trained bureau staff who used their judgment to match and code records. The four phases of the person matching process were (1) computer matching, (2) clerical matching, (3) nationwide field follow-up on records requiring more information, and (4) a second phase of clerical matching after field follow-up.[1] Each subsequent phase used additional information and matching rules in an attempt to match records that the previous phase could not link.

Computer Matching



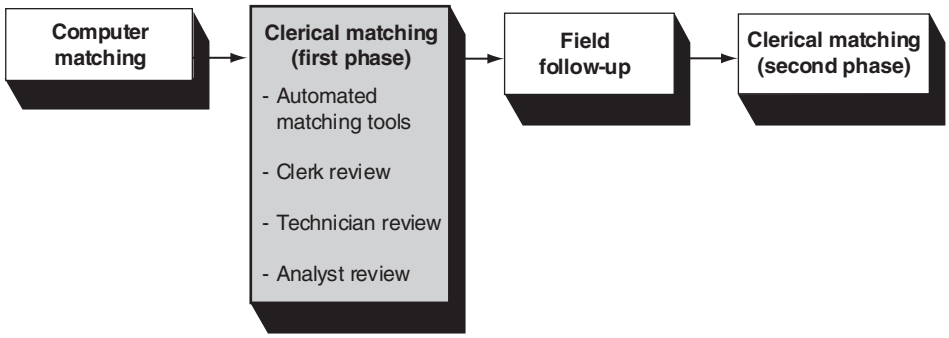
Computer matching took pairs of census and A.C.E. records and compared various personal characteristics such as name, age, and gender. The computer then calculated a match score for the paired records based on the extent to which the personal characteristics were aligned. Experienced bureau staff reviewed the lists of paired records, sorted by their match

[1]A person record should have contained the following characteristics: first name, last name, middle name, gender, race, Hispanic origin, age, date of birth, and relationship to the respondent of the A.C.E. or the census.

scores, and judgmentally assigned cutoff scores. The cutoff scores were break points used to categorize the paired records into one of three groups so that the records could be coded as a "match," "possible match," or one of a number of codes that defines them as not matched. Computer matching successfully assigned a match score to nearly 1 million of the more than 1.4 million records reviewed (about 66 percent).

Bureau staff documented the cutoff scores for each of the match groups. However, they did not document the criteria or rules used to determine cutoff scores, the logic of how they applied them, and examples of their application. As a result, the bureau may not benefit from the possible lessons learned on how to apply cutoff scores. When the computer links few records as possible matches, clerks will spend more time searching records and linking them. In contrast, when the computer links many records as possible matches, clerks will spend less time searching for records to link and more time unlinking them. Without documentation and knowledge of the effect of cutoff scores on clerical matching productivity, future bureau staff will be less able to determine whether to set cutoff scores to link few or many records together as possible matches.

First Phase of Clerical Matching

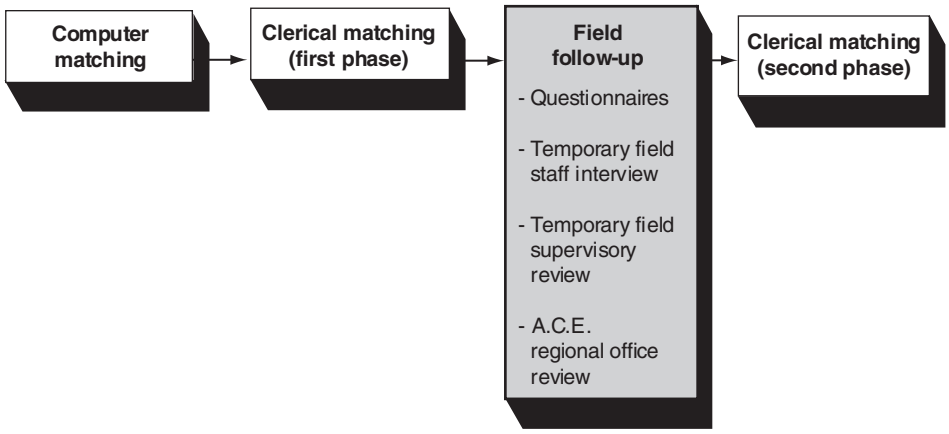


During clerical matching, three levels of matchers—including over 200 clerks, about 40 technicians, and 10 experienced analysts or "expert matchers"—applied their expertise and judgment to manually match and code records. A computer software system managed the workflow of the clerical matching stages. The system also provided access to additional information, such as electronic images of census questionnaires that could assist matchers in applying criteria to match records. According to a bureau official, a benefit of clerical matching was that records of entire households could be reviewed together, rather than just individually as in

computer matching. During this phase over a quarter million records (or about 19 percent) were assigned a final match code.

The bureau taught clerks how to code records in situations in which the A.C.E. and census records differed because one record contained a nickname and the other contained the birth name. The bureau also taught clerks how to code records with abbreviations, spelling differences, middle names used as first names, and first and last names reversed. These criteria were well documented in both the bureau's procedures and operations memorandums and clerical matchers' training materials, but how the criteria were applied depended on the judgment of the matchers. The bureau trained clerks and technicians for this complex work using as examples some of the most challenging records from the 1998 Dress Rehearsal person matching operation. In addition, the analysts had extensive matching experience. For example, the 4 analysts that we interviewed had an average of 10 years of matching experience on other decennial census surveys and were directly involved in developing the training materials for the technicians and clerks.

Field Follow-up

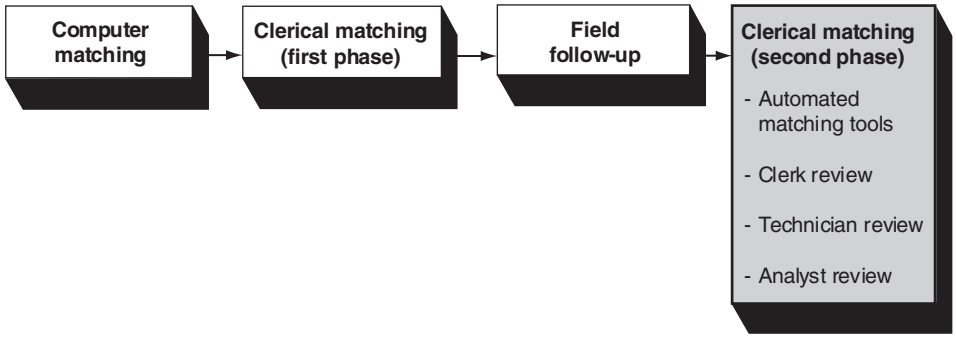


The bureau conducted a nationwide field follow-up on over 213,000 records (or about 15 percent) for which the bureau needed additional information before it could accurately assign a match code. For example, sometimes matchers needed additional information to verify that possibly matched records were actually records of the same person, that a housing unit was located in the sample area on Census Day, or that a person lived in the sample area on Census Day. Field follow-up questionnaires were printed at

the National Processing Center and sent to the appropriate A.C.E. regional office.

Field follow-up interviewers from the bureau's regional offices were required to visit specified housing units and obtain information from a knowledgeable respondent. If the household member for the record in question still lived at the A.C.E. address at the time of the interview and was not available to be interviewed after six attempts, field follow-up interviewers were allowed to obtain information from one or more knowledgeable proxy respondents, such as a landlord or neighbor.

Second Phase of Clerical Matching



The second phase of clerical matching used the information obtained during field follow-up in an attempt to assign a final match code to records. As in the first phase of clerical matching, the criteria used to match and code records were well documented in both the bureau's procedures and operations memorandums and clerical matchers' training materials. Nevertheless, in applying those criteria, clerical matchers had to use their own judgment and expertise. This was particularly true when matching records that contained incomplete and inconsistent information, as noted in the following examples.

- Different household members provided conflicting information.

The census counted one person—the field follow-up respondent. A.C.E. recorded four persons—including the respondent and her daughter. The respondent, during field follow-up, reported that all four persons recorded by A.C.E. lived at the housing unit on Census Day. During the field follow-up interview, the respondent's daughter came to the house and disagreed with the respondent. The interviewer changed the answers

on the field follow-up questionnaire to reflect what the daughter said—the respondent was the only person living at the household address on Census Day. The other three people were coded as not living at the household address on Census Day. According to bureau staff, the daughter's response seemed more reliable.

- An interviewer's notes on the field follow-up questionnaire conflicted with recorded information.

The census counted 13 people—including the respondent and 2 people not matched to A.C.E. records. A.C.E. recorded 12 people—including the respondent, 10 other matched people, and the respondent's daughter who was not matched to census records. The field follow-up interview attempted to resolve the unmatched census and A.C.E. people. Answers to questions on the field follow-up questionnaire verified that the daughter lived at the housing address on Census Day. However, the interviewer's notes indicated that the daughter and the respondent were living in a shelter on Census Day. The daughter was coded as not living at the household address on Census Day, while the respondent remained coded as matched and living at the household address on Census Day. According to bureau staff, the respondent should also have been coded as a person that did not live at the household address on Census Day, based on the notes on the field follow-up questionnaire.

- A.C.E., census, or both counted people at the wrong address.

The census counted two people—the respondent and her husband—twice; once in an apartment and once in a business office that the husband worked in, both in the same apartment building. The A.C.E. did not record anyone at either location, as the residential apartment was not in the A.C.E. interview sample. The respondent, during field follow-up, reported that they lived at their apartment on Census Day and not at the business office. The couple had responded to the census on a questionnaire delivered to the business office. A census enumerator, following up on the "nonresponse" from the couple's apartment, had obtained census information from a neighbor about the couple. The couple, as recorded by the census at the business office address, was coded as correctly counted in the census. The couple, as recorded by the census at the apartment address, was coded as living outside the sample block. According to bureau staff, the couple recorded at the business office address were correctly coded, but the couple recorded at the apartment should have been coded as duplicates.

- An uncooperative household respondent provided partial or no information.

The census counted a family of four—the respondent, his wife, and two daughters. A.C.E. recorded a family of three—the same husband and wife, but a different daughter's name, "Buffy." The field follow-up interview covered the unmatched daughters—two from census and one from A.C.E. The respondent confirmed that the four people counted by the census were his family and that "Buffy" was a nickname for one of his two daughters, but he would not identify which one. The interviewer wrote in the notes that the respondent "was upset with the number of visits" to his house. "Buffy" was coded as a match to one of the daughters; the other daughter was coded as counted in the census but missed by A.C.E. According to bureau staff, since the respondent confirmed that "Buffy" was a match for one of his daughters—although not which one—and that four people lived at the household address on Census Day, they did not want one of the daughters coded so that she was possibly counted as a missed census person.

Since each record had to have a code identifying whether it was a match by the end of the second clerical matching phase, records that did not contain enough information after field follow-up to be assigned any other code were coded as "unresolved." The bureau later imputed the match code results for these records using statistical methods. While imputation for some situations may be unavoidable, it introduces uncertainty into estimates of census over- or undercount rates. The following are examples of situations that resulted in records coded as "unresolved."

- Conflicting information was provided for the same household.

The census counted four people—a woman, an "unmarried partner," and two children. A.C.E. recorded three people—the same woman and two children. During field follow-up, the woman reported to the field follow-up interviewer that the "unmarried partner" did not really live at the household address, but just came around to baby-sit, and that she did not know where he lived on Census Day. According to bureau staff, probing questions during field follow-up determined that the "unmarried partner" should not have been coded as living at the housing unit on Census Day. Therefore, the "unmarried partner" was coded as "unresolved."

- A proxy respondent provided conflicting or inaccurate information.

The census counted one person—a female renter. A.C.E. did not record anyone. The apartment building manager, who was interviewed during field follow-up, reported that the woman had moved out of the household address sometime in February 2000, but the manager did not know the woman's Census Day address. The same manager had responded to an enumerator questionnaire for the census in June 2000 and had reported that the woman did live at the household address on Census Day. The woman was coded as "unresolved."

Quality Assurance Results Suggest Person Matching Procedures Were Implemented as Planned

The bureau employed a series of quality assurance procedures for each phase of person matching. The bureau reported that person matching quality assurance was successful at minimizing errors because the quality assurance procedures found error rates of less than 1 percent.

Computer Matching

Clerks were to review all of the match results to ensure, among other things, that the records linked by the computer were not duplicates and contained valid and complete names. Moreover, according to bureau officials, the software used to link records had proven itself during a similar operation conducted for the 1990 Census. The bureau did not report separately on the quality of computer matched records. Although there were no formal quality assurance results from computer matching, at our request the bureau tabulated the number of records that the computer had coded as "matched" that had subsequently been coded otherwise. According to the bureau, the subsequent matching process resulted in a different match code for about 0.6 percent of the almost 500,000 records initially coded as matched by the computer. Of those records having their codes changed by later matching phases, over half were eventually coded as duplicates and almost all of the remainder were rematched to someone else.

Two Phases of Clerical Matching

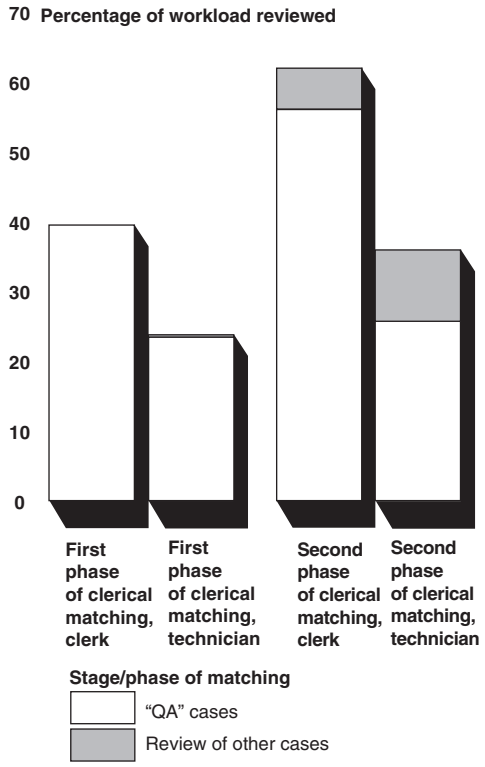
Technicians reviewed the work of clerks and analysts reviewed the work of technicians primarily to find clerical errors that (1) would have prevented records from being sent to field follow-up, (2) could cause a record to be incorrectly coded as either properly or erroneously counted by the census,

or (3) would cause a record to be incorrectly removed from the A.C.E. sample. Analysts' work was not reviewed.

Clerks and technicians with error rates of less than 4 percent had a random sample of about 25 percent of their work reviewed, while clerks and technicians exceeding the error threshold had 100 percent of their work reviewed. About 98 percent of clerks in the first phase of matching had only a sample of their work reviewed. According to bureau data, less than 1 percent of match decisions were revised during quality assurance reviews, leading the bureau to conclude that clerical matching quality assurance was successful.

Under certain circumstances, technicians and analysts performed additional reviews of clerks' and technicians' work. For example, if during the first phase of clerical matching a technician had reviewed and changed more than half of a clerk's match codes in a given geographic cluster, the cluster was flagged for an analyst to review all of the clerk and technician coding for that area. During the second phase, analysts were required to make similar reviews when only one of the records was flagged for their review. This is one of the reasons why, as illustrated in figure 2, these additional reviews were a much more substantial part of the clerks' and technicians' workload that was subsequently reviewed by more senior matchers. The total percentage of workload reviewed ranged from about 20 to 60 percent across phases of clerical matching, far in excess of the 11-percent quality assurance level for the bureau's person interviewing operation.

Figure 2: Person Matching, Quality Assurance Coverage



Source: GAO analysis of U.S. Census Bureau data.

Field Follow-up

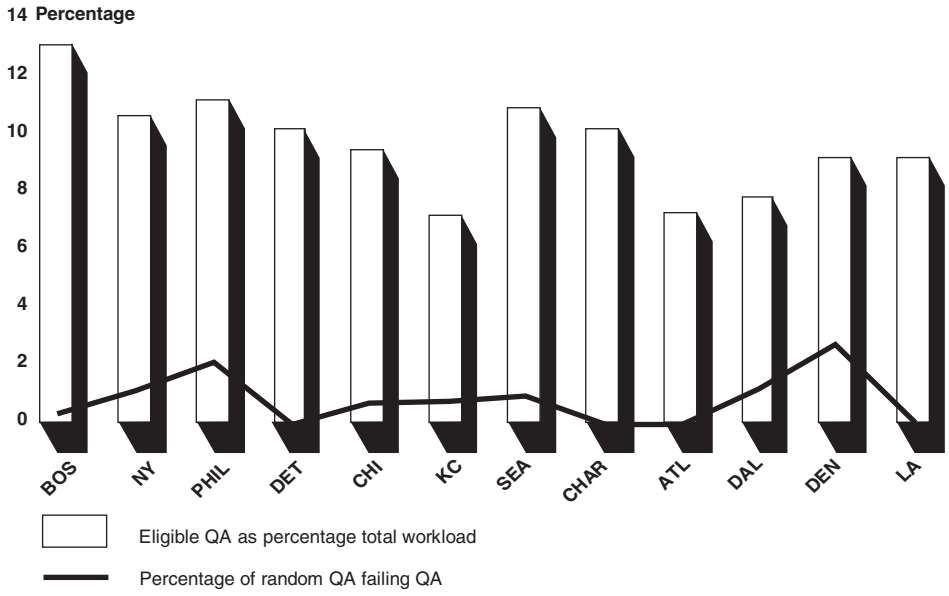
The quality assurance plan for the field follow-up phase had two general purposes: (1) to ensure that questionnaires had been completed properly and legibly and (2) to detect falsification.[2] Supervisors initially reviewed each questionnaire for legibility and completeness. These reviews also checked the responses for consistency. Office staff were to conduct similar reviews of each questionnaire.

To detect falsification, the bureau was to review and edit each questionnaire at least twice and recontact a random sample of 5 percent of

[2]According to the bureau, a questionnaire failed quality assurance if a respondent said that the original follow-up interviewer did not contact him or her for the original interview.

the respondents. As shown in figure 3, all 12 of the A.C.E. regional offices exceeded the 5 percent requirement by selecting more than 7 percent of their workload for quality assurance review, and the national rate of quality assurance review was about 10 percent.

Figure 3: Quality Assurance of Field Follow-up by A.C.E. Regional Office



Source: GAO analysis of U.S. Census Bureau data.

At the local level, however, there was greater variation. There are many reasons why the quality assurance coverage can appear to vary locally. For example, a local census area could have a low quality assurance coverage rate because interviewers in that area had their work reviewed in other areas, or the area could have had an extremely small field follow-up workload, making the difference of just one quality assurance questionnaire constitute a large percentage of the local workload. Seventeen local census office areas (out of 520 nationally, including Puerto Rico) had 20 percent or more of field follow-up interviews covered by the quality assurance program, and, at the other extreme, 5 local census areas had 5 percent or less of the work covered by the quality assurance program. Less than 1 percent of the randomly selected questionnaires failed quality assurance nationally, leading the bureau to report this quality assurance operation as successful.

When recontacting respondents to detect falsification by interviewers, quality assurance supervisors were to determine whether the household had been contacted by an interviewer, and if it had not, the record of that household failed quality assurance. According to bureau data, about 0.8 percent of the randomly selected quality assurance questionnaires failed quality assurance nationally. This percentage varied between 0 and about 3 percent across regions.

The Bureau Took Action to Address Some Deviations, but Effect on Matching Results Is Unknown

The bureau carried out person matching as planned, with only a few procedural deviations. Although the bureau took action to address these deviations, it has not determined how matching results were affected. As shown in table 1, these deviations included (1) census files that were delivered late, (2) a programming error in the clerical matching software, (3) printing errors in field follow-up forms, (4) regional offices that sent back incomplete questionnaires, and (5) the need for additional time to complete the second phase of clerical matching.

It is unknown what, if any, cumulative effect these procedural deviations may have had on the quality of matching for these records or on the resultant A.C.E. estimates of census undercounts. However, bureau officials believe that the effect of the deviations was small based on the timely responses taken to address them. The bureau conducted reinterviewing and re-matching studies on samples of the 2000 A.C.E. sample and concluded that matching quality in 2000 was improved over that in 1990, but that error introduced during matching operations remained and contributed to an overstatement of A.C.E. estimates of the census undercounts. The studies provided some categorical descriptions of the types of matching errors measured, but did not identify the procedural causes, if any, for those errors. Furthermore, despite the improvement in matching reported by the bureau, A.C.E. results were not used to adjust the census due to these errors as well as other remaining uncertainties. The bureau has reported that additional review and analysis on these remaining uncertainties would be necessary before any potential uses of these data can be considered.

Table 1: Deviations from the Planned Person Matching Operation

Deviation	Corrective action taken	Effect on process
Late delivery of census files.	Bureau employees worked extra hours to make up the time.	Computer matching was started 3 days later than scheduled and finished 1 day behind schedule.
Programming error in clerical matching software.	The number of records to be completed between error rate calculations was modified twice in the software managing the quality assurance of clerical matching and the software problem was quickly fixed.	Assignments of sampled or 100-percent review of clerks' and technicians' work were made manually for 2 days.
1. Programming error caused errors in printing last names. 2. Other printing problems.	1. Printing of field follow-up questionnaires was suspended temporarily. The procedure was supplemented. 2. No action taken because bureau staff viewed it as insignificant.	1. Extra steps were taken during matching for 5 percent of records. This slowed each region's questionnaire processing for 1 to 4 days. 2. The effect is unknown, but bureau staff viewed it as insignificant.
Regional offices sent back incomplete field follow-up questionnaires that contained a section to verify whether a housing unit was in the A.C.E. sample.	Forty-eight incomplete field follow-up questionnaires were returned to the regional offices during the first 6 days of the second clerical matching phase.	The effect is unknown because the total number of questionnaires with this section incomplete is not known.
Extra time was needed to complete the second phase of clerical matching.	The schedule for the second phase of clerical matching was extended.	Subsequent A.C.E. operations had to make up the time.

Late Delivery of Census Files Delayed Computer Matching Start

The computer matching phase started 3 days later than scheduled and finished 1 day late due to the delayed delivery of census files. In response, bureau employees who conducted computer matching worked overtime hours to make up lost time. Furthermore, A.C.E. regional offices did not receive clusters in the prioritized order that they had requested. The reason for prioritizing the clusters was to provide as much time as possible for field follow-up on clusters in the most difficult areas. Examples of areas that were expected to need extra time were those with staffing difficulties, larger workloads, or expected weather problems. Based on the bureau's Master Activities Schedule, the delay did not affect the schedule of

subsequent matching phases. Also, bureau officials stated that although clusters were not received in prioritized order, field follow-up was not greatly affected because the first clerical matching phase was well staffed and sent the work to regional offices quickly.

Programming Error and Analyst Backlog Required Software Modifications during Clerical Matching

On the first full day of clerical matching, the bureau identified a programming error in the quality assurance management system, which made some clerks and technicians who had not passed quality assurance reviews appear to have passed. In response, bureau officials manually overrode the system. Bureau officials said the programming error was fixed within a couple of days, but could not explain how the programming error occurred. They stated that the software system used for clerical matching was thoroughly tested, although it was not used in any prior censuses or census tests, including the Dress Rehearsal. As we have previously noted, programming errors that occur during the operation of a system raise questions about the development and acquisition processes used for that system.[3]

Field Follow-up Questionnaires Contained Printing Errors

A programming error caused last names to be printed improperly on field follow-up forms for some households containing multiple last names. In situations in which regional office staff may not have caught the printing error and interviewers may have been unaware of the error—such as when those questionnaires were completed before the problem was discovered—interviews may have been conducted using the wrong last name, thus recording misleading information. According to bureau officials, in response, the bureau (1) stopped printing questionnaires on the date officials were notified about the misprinted questionnaires, (2) provided information to regional offices that listed all field follow-up housing units with multiple names that had been printed prior to the date the problem was resolved, and (3) developed procedures for clerical matchers to address any affected questionnaires being returned that had not been corrected by regional office staff. While resolving the problem, productivity was initially slowed in the A.C.E. regional offices for approximately 1 to 4 days, yet field follow-up was completed on time.

[3]U.S. General Accounting Office, *2000 Census: Headquarters Processing System Status and Risks*, GAO-01-1 (Washington, D.C.: October 17, 2000).

Bureau officials inadvertently introduced this error when they addressed a separate programming problem in the software. Bureau officials stated that they tested this software system; however, the system was not given a trial run during the Census Dress Rehearsal in 1998. According to bureau officials, the problem did not affect data quality because it was caught early in the operation and follow-up forms were edited by regional staff. However, the bureau could not determine the exact day of printing for each questionnaire and thus did not know exactly which households had been affected by the problem. According to bureau data, the problem could have potentially affected over 56,000 persons, or about 5 percent of the A.C.E. sample.

In addition to the problem printing last names, the bureau experienced other printing problems. According to bureau staff, field follow-up received printed questionnaires that were (1) missing pages, (2) missing reference notes written by clerical matchers, and (3) missing names and/or having some names printed more than once for some households of about nine or more people. According to bureau officials, these problems were not resolved during the operation because they were reported after field follow-up had started and the bureau was constrained by deadlines. Bureau officials stated that they believed that these problems would not significantly affect the quality of data collected or match code results, although bureau officials were unable to provide data that would document either the extent, effect, or cause of these problems.

Regional Offices Sent Back Incomplete Field Follow-up Questionnaires

The bureau's regional offices submitted questionnaires containing an incomplete "geocoding" section. This section was to be used in instances when the bureau needed to verify whether a housing unit (1) existed on Census Day and (2) was correctly located in the A.C.E. sample area. Although the bureau returned 48 questionnaires during the first 6 days of the operation to the regional offices for completion, bureau officials stated that after that they no longer returned questionnaires to the regional offices because they did not want to delay the completion of field follow-up.

A total of over 10,000 questionnaires with "geocoding" sections were initially sent to the regional offices. The bureau did not have data on the number, if any, of questionnaires that the regional offices submitted incomplete beyond the initial 48. The bureau would have coded as "unresolved" the persons covered by any incomplete questionnaires. As previously stated, the bureau later imputed the match code results for

these records using statistical methods, which could introduce uncertainty into estimates of census over- or undercount rates.

According to bureau officials, this problem was caused by (1) not printing a checklist of all sections that needed to be completed by interviewers, (2) no link from any other section of the questionnaire to refer interviewers to the "geocoding" section, and (3) field supervisors following the same instructions as interviewers to complete their reviews of field follow-up forms. However, bureau officials believed that the mistake should have been caught by regional office reviews before the questionnaires were sent back for processing.

Extra Time Was Needed to Complete the Second Phase of Clerical Matching

About a week after the second clerical matching phase began, officials requested an extension, which was granted for 5 days, to complete the second clerical matching phase. According to bureau officials, the operation could have been completed by the November 30, 2000, deadline as planned, but they decided to take extra steps to improve data quality that required additional time. According to bureau officials, the delay in completing person matching had no effect on the final completion schedule, only the start of subsequent A.C.E. processing operations.

Conclusions

Matching A.C.E. and census records was an inherently complex and labor-intensive process that often relied on the judgment of trained staff, and the bureau prepared itself accordingly. For example, the bureau provided extensive training for its clerical matchers, generally provided thorough documentation of the process and criteria to be used in carrying out their work, and developed quality assurance procedures to cover its critical matching operations. As a result, our review identified few significant operational or procedural deviations from what the bureau planned, and the bureau took timely action to address them.

Nevertheless, our work identified opportunities for improvement. These opportunities include a lack of written documentation showing how cutoff scores were determined and programming errors in the clerical matching software and software used to print field follow-up forms. Without written documentation, the bureau will be less likely to capture lessons learned on how cutoff scores should be applied, in order to determine the impact on clerical matching productivity. Moreover, the discovery of programming errors so late in the operation raises questions about the development and

acquisition processes used for the affected A.C.E. computer systems. In addition, one lapse in procedures may have resulted in incomplete geocoding sections verifying that the person being matched was in the geographic sample area. The collective effect that these deviations may have had on the accuracy of A.C.E. results is unknown. Although the bureau has concluded that A.C.E. matching quality improved compared to 1990, the bureau has reported that error introduced during matching operations remained and contributed to an overstatement of the A.C.E. estimate of census undercounts. To the extent that the bureau employs an operation similar to A.C.E. to measure the quality of the 2010 Census, it will be important for the bureau to determine the impact of the deviations and explore operational improvements, in addition to the research it might carry out on other uncertainties in the A.C.E. results.

Recommendations for Executive Action

As the bureau documents its lessons learned from the 2000 Census and continues its planning efforts for 2010, we recommend that the secretary of commerce direct the bureau to take the following actions:

1. Document the criteria and the logic that bureau staff used during computer matching to determine the cutoff scores for matched, possibly matched, and unmatched record pairs.

2. Examine the bureau's system development and acquisition processes to determine why the problems with A.C.E. computer systems were not discovered prior to deployment of these systems.

3. Determine the effect that the printing problems may have had on the quality of data collected for affected records, and thus the accuracy of A.C.E. estimates of the population.

4. Determine the effect that the incomplete geocoding section of the questionnaires may have had on the quality of data collected for affected records, and thus the accuracy of A.C.E. estimates of census undercounts.

Agency Comments and Our Evaluation

The secretary of commerce forwarded written comments from the U.S. Census Bureau on a draft of this report. (See appendix II.) The bureau had no comments on the text of the report and agreed with, and is taking action on, two of our four recommendations.

In responding to our recommendation to document the criteria and the logic that bureau staff used during computer matching to determine cutoff scores, the bureau acknowledged that such documentation may be informative and that such documentation is under preparation. We look forward to reviewing the documentation when it is complete.

In responding to our recommendation to examine system development and acquisition processes to determine why problems with the A.C.E. computer systems were not discovered prior to deployment, the bureau responded that despite extensive testing of A.C.E. computer systems, a few problems may remain undetected. The bureau plans to review the process to avoid such problems in 2010, and we look forward to reviewing the results of their review.

Finally, in response to our two recommendations to determine the effects that printing problems and incomplete questionnaires had on the quality of data collected and the accuracy of A.C.E. estimates, the bureau responded that it did not track the occurrence of these problems because the effects on the coding process and accuracy were considered to be minimal since all problems were identified early and corrective procedures were effectively implemented. In our draft report we recognized that the bureau took timely corrective action in response to these and other problems that arose during person matching. Yet we also reported that bureau studies of the 2000 matching process had concluded that matching error contributed to error in A.C.E. estimates without identifying procedural causes, if any. Again, to the extent that the bureau employs an operation similar to A.C.E. to measure the quality of the 2010 Census, it will be important for the bureau to determine the impact of the problems and explore operational improvements as we recommend.

We are sending copies of this report to other interested congressional committees. Please contact me on (202) 512-6806 if you have any questions. Key contributors to this report are included in appendix III.

Patricia A. Dalton
Director
Strategic Issues

Scope and Methodology

To address our three objectives, we examined relevant bureau program specifications, training manuals, office manuals, memorandums, and other progress and research documents. We also interviewed bureau officials at bureau headquarters in Suitland, Md., and the bureau's National Processing Center in Jeffersonville, Ind., which was responsible for the planning and implementation of the person matching operation.

In addition, to review the process and criteria involved in making an A.C.E. and census person match, we observed the match clerk training at the National Processing Center and a field follow-up interviewer training session in Dallas, Tex. To identify the results of the quality assurance procedures used in key person matching phases, we analyzed operational data and reports provided to us by the bureau, as well as extracts from the bureau's management information system, which tracked the progress of quality assurance procedures. Other independent sources of the data were not available for us to use to test the data that we extracted, although we were able to corroborate data results with subsequent interviews of key staff.

Finally, to examine how, if at all, the matching operation deviated from what was planned, we selected 11 locations in 7 of the 12 bureau census regions (Atlanta, Chicago, Dallas, Denver, Los Angeles, New York, and Seattle).[4] At each location we interviewed A.C.E. workers from November through December 2000. The locations selected for field visits were chosen primarily for their geographic dispersion (i.e., urban or rural), variation in type of enumeration area (e.g., update/leave or list enumerate), and the progress of their field follow-up work. In addition, we reviewed the match code results and field follow-up questionnaires from 48 sample clusters. These clusters were chosen because they corresponded to the local census areas we visited and contained records reviewed during every phase of the person matching operation. The results of our field visits and our cluster review are not generalizable nationally to the person matching operation.

We performed our audit work from September 2000 through September 2001 in accordance with generally accepted government auditing standards.

[4]The 11 locations we visited were Chicago, Ill.; Miami and Lakeland, Fla.; New York, N.Y.; McAllen, Beaumont, and Houston, Tex.; Los Angeles, Calif.; Seattle, Wash.; and Phoenix and Window Rock, Ariz.

Comments from the Department of Commerce

THE SECRETARY OF COMMERCE
Washington, D.C. 20230

FEB 13 2002

Mr. J. Christopher Mihm
Director, Strategic Issues
U.S. General Accounting Office
Washington, DC 20548

Dear Mr. Mihm:

The Department of Commerce appreciates the opportunity to comment on the General Accounting Office (GAO) draft document entitled *2000 Census: Coverage Evaluation Matching Implemented As Planned, but Census Bureau Should Evaluate Lessons Learned.* The Department has no comment on the text of this report. Our responses to GAO's recommendations are enclosed.

Warm regards,

Donald L. Evans

Enclosure

Comments from the U.S. Department of Commerce
U.S. Census Bureau

U.S. General Accounting Office draft report entitled *2000 Census: Coverage Evaluation Matching Implemented As Planned, but Census Bureau Should Evaluate Lessons Learned*

Comments on the Text of the Report

The U.S. Census Bureau has no comments on the text of the report.

Responses to GAO Recommendations

1. Document the criteria and the logic that Bureau staff used during computer matching to determine the cutoff scores for matched, possibly matched, and unmatched record pairs.

Census Bureau Response: *The Census Bureau has acknowledged that such a document may be informative. As such, a document is under preparation.*

2. Examine the Bureau's system development and acquisition processes to determine why the problems with A.C.E. computer systems were not discovered prior to deployment of these systems.

Census Bureau Response: *The Census Bureau conducted extensive systematic testing on the A.C.E. computer systems; however, these systems are inherently complex, and a few problems may have remained undetected in spite of extensive testing. The problem identified in the report was primarily related to a software program that was developed after the Dress Rehearsal, and its testing was not as extensive as what was done for the other components of the system. We plan to review our system development processes to avoid similar problems in the 2010 census.*

3. Determine the effect that the printing problems may have had on the quality of data collected for affected records, and thus the accuracy of A.C.E. estimates of the population.

Census Bureau Response: *When the printing problems were identified, it was thought that they would not significantly affect the coding process; therefore, we did not track the incidence of the problems and cannot report on the effect of these problems. However, the effect on the accuracy of the A.C.E. estimates is believed to be minimal, because the problems were identified early and corrective procedures were effectively implemented.*

2

4. Determine the effect that the incomplete geocoding section of the questionnaires may have had on the quality of data collected for affected records, and thus the accuracy of A.C.E. estimates of census undercounts.

Census Bureau Response: *As in Item 3, we did not track the incidence of such cases, because the effects on accuracy were believed to be minimal, given that the problem was identified early and corrective procedures were effectively implemented.*

GAO Contact and Staff Acknowledgments

GAO Contact

Robert Goldenkoff, (202) 512-2757

Acknowledgments

In addition to those named above, Ty Mitchell, Lynn Wasielewski, Steven Boyles, Angela Pun, J. Christopher Mihm, and Richard Hung contributed to this report.

GAO's Mission

The General Accounting Office, the investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents is through the Internet. GAO's Web site (www.gao.gov) contains abstracts and full-text files of current reports and testimony and an expanding archive of older products. The Web site features a search engine to help you locate documents using key words and phrases. You can print these documents in their entirety, including charts and other graphics.

Each day, GAO issues a list of newly released reports, testimony, and correspondence. GAO posts this list, known as "Today's Reports," on its Web site daily. The list contains links to the full-text document files. To have GAO E-mail this list to you every afternoon, go to www.gao.gov and select "Subscribe to daily e-mail alert for newly released products" under the GAO Reports heading.

Order by Mail or Phone

The first copy of each printed report is free. Additional copies are $2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to:

U.S. General Accounting Office
P.O. Box 37050
Washington, D.C. 20013

To order by Phone: Voice: (202) 512-6000
TDD: (202) 512-2537
Fax: (202) 512-6061

Visit GAO's Document Distribution Center

GAO Building
Room 1100, 700 4th Street, NW (corner of 4th and G Streets, NW)
Washington, D.C. 20013

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm,
E-mail: fraudnet@gao.gov, or
1-800-424-5454 or (202) 512-7470 (automated answering system).

Public Affairs

Jeff Nelligan, Managing Director, NelliganJ@gao.gov (202) 512-4800
U.S. General Accounting Office, 441 G. Street NW, Room 7149,
Washington, D.C. 20548

PRINTED ON RECYCLED PAPER

United States
General Accounting Office
Washington, D.C. 20548-0001

Official Business
Penalty for Private Use $300

Address Correction Requested

Presorted Standard
Postage & Fees Paid
GAO
Permit No. GI00