

December 2001

2000 CENSUS

Coverage Evaluation Interviewing Overcame Challenges, but Further Research Needed



Contents

Letter		1
	Background	2
	Results in Brief	4
	Scope and Methodology	7
	Conclusions	22
	Recommendations for Executive Action	24
	Agency Comments and Our Evaluation	24

Appendixes

Appendix I: Comments from the Secretary of Commerce	28
GAO Comments	35
Appendix II: GAO Contacts and Staff Acknowledgments	41
GAO Contacts	41
Acknowledgments	41

Tables

Table 1: Deliverability of Initial Address List for Major Census Operations	20
---	----

Figures

Figure 1: A.C.E. Survey Followed Steps Similar to Census	3
Figure 2: Proportion of Person Interviewing Workload Completed by Telephone Interviews by Census Region	10
Figure 3: Distribution of 520 Local Census Office Areas' Share of Person Interviewing Completed During Nonresponse Conversion	11
Figure 4: Person Interviewing Quality Assurance by Census Region	14
Figure 5: Distribution of Proxy Rates for 520 Local Census Office Areas	17
Figure 6: Distribution of 520 Local Census Office Partial Interview Rates	18

Contents



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

December 31, 2001

The Honorable Joseph I. Lieberman
Chairman
The Honorable Fred Thompson
Ranking Minority Member
Committee on Governmental Affairs
United States Senate

The Honorable Dan Burton
Chairman
The Honorable Henry A. Waxman
Ranking Minority Member
Committee on Government Reform
House of Representatives

The Honorable Dan Miller
Chairman
The Honorable William Lacy Clay, Jr.
Ranking Minority Member
Subcommittee on the Census
Committee on Government Reform
House of Representatives

The Honorable Carolyn B. Maloney
House of Representatives

This report provides you with the results of our review of the U.S. Census Bureau's Accuracy and Coverage Evaluation (A.C.E.) person interviewing operation. A.C.E. used an independent sample survey to assess the quality of the population data collected by the 2000 Census by estimating the number of people missed, counted more than once, or otherwise improperly counted in the census. Partly on the basis of the A.C.E. results, the Acting Director of the Census Bureau recommended on March 1, 2001, that the 2000 Census tabulations for purposes of redrawing the boundaries of congressional districts not be adjusted, and on October 16, 2001, he similarly recommended that unadjusted census data be used for nonredistricting purposes. These decisions will have far-ranging implications because census data are used to distribute billions of dollars in federal funding, guide public and private investment decisions, and provide a baseline for a number of other statistical measurement programs.

In addition, the results of A.C.E. are expected to play an important role in the Bureau's research and preparation for the 2010 Census.

Person interviewing was a critical component of A.C.E. because it was used to collect the sample data used to evaluate the nationwide headcount. It was conducted from April 24 through September 11, 2000. In our prior work, we noted that the Bureau's plans for assessing the quality of census population data faced several methodological, technological, and quality control challenges. This report is the latest in our series of reviews that examine the results of key census-taking operations and highlight opportunities for reform. (See app. IV for a list of products issued to date.) As agreed, our objective for this report was to identify the challenges, if any, that the Bureau confronted during person interviewing and the degree to which the Bureau successfully addressed any challenges.

Background

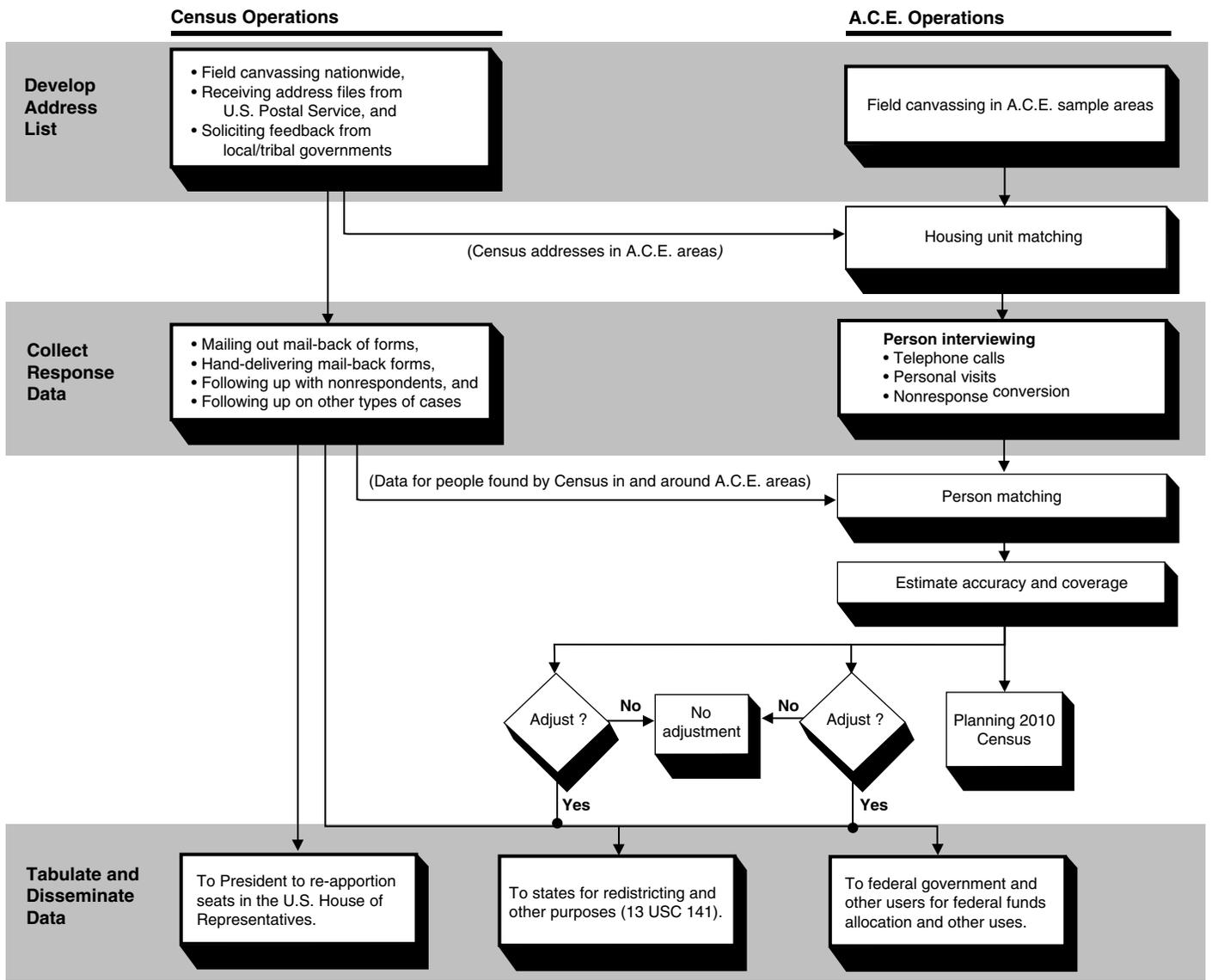
The Census Bureau conducted A.C.E. on a sample of areas across the country to estimate the number of people and housing units missed or counted more than once in the census and to evaluate the final census counts.

The statistical methodology underpinning A.C.E. assumes that the chance that a person is counted by the census is not affected by whether he or she is counted in A.C.E., or vice versa. Violating this "independence" assumption can bias the estimate of the number of people missed in the census and thus either overstate or understate the census undercount. The Bureau's procedures called for it to go to great lengths to maintain this independence. As illustrated in figure 1, the Bureau developed separate address lists—one for the entire nation of about 120 million housing units and one for A.C.E. sample areas—and collected data through two independent operations. For the census, the Bureau mailed out forms for mail-back to most of the housing units in the country; hand-delivered mail-back forms to most of the rest of the country (in an operation called Update/Leave); and then carried out a number of other follow-up operations, the largest of which was called nonresponse follow-up. A.C.E. collected its response data during person interviewing from April 24 through September 11, 2000, with telephone calls or visits to about 314,000 housing units.

A.C.E. person interviewing was managed directly out of 12 A.C.E. regional offices, independent of the 12 regional census centers from which the census was managed. A.C.E. regional offices managed person interviewing

workflow at the geographic level of the local census office area out of convenience. There were 520 local census offices operating during the census, including 9 in Puerto Rico managed out of the Boston regional office, and person interviewing took place in the area of each.

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



Source: U.S. Census Bureau documents.

The person interviewing operation had three phases: early telephone interviewing; then interviewing conducted by personal visits; and finally nonresponse conversion, when difficult cases were reassigned to the operation's best interviewers to reduce the number of noninterviews. In each phase, interviewers relied on an automated survey instrument and databases stored on laptop computers assigned to each interviewer. By having the interviewers use laptops to dial in to the Bureau's servers, the Bureau could manage cases automatically and remotely.

In its initial design for the 2000 Census, the Bureau planned a "one-number" census that would have integrated the results of a survey similar to A.C.E. with the traditional census to provide one adjusted set of numbers by December 31, 2000. However, the U.S. Supreme Court ruled in January 1999 that statistical sampling could not be used to generate population data for reapportioning the House of Representatives.¹ Following that ruling, the Bureau abandoned its plans to conduct a one-number census using statistical methods integrated into the final census counts.

On December 28, 2000, the Bureau delivered its population counts for purposes of reapportioning seats in the House of Representatives to the Secretary of Commerce for his transmission to the President. On March 1, 2001, a committee of senior Bureau executives recommended that unadjusted census data be released as the Bureau's official redistricting data. The Acting Director of the U.S. Census Bureau concurred, and the Secretary of Commerce announced on March 7, 2001, his decision to release the unadjusted data. On October 16, 2001, the Acting Director of the Census Bureau decided that unadjusted data should be used for nonredistricting purposes as well as for postcensus population estimates and for benchmarks for other federal surveys. The Bureau is continuing to investigate issues related to A.C.E. and the Census, and the results of that investigation are expected to influence the Bureau's planning for the 2010 Census.

Results in Brief

In conducting person interviewing, the Bureau faced a number of operational challenges, including (1) completing the operation on schedule, (2) ensuring data quality, (3) dealing with unexpected computer problems, (4) obtaining a quality address list, and (5) keeping person interviewing

¹*Department of Commerce vs. United States House of Representatives* 525 U.S. 316 (1999).

independent of census follow-up operations as necessary for unbiased estimates of census errors.

The Bureau completed the person interviewing operation largely ahead of schedule. Timeliness was important because the Bureau believes that data quality declines the longer data collection continues, and subsequent data processing required that data be collected on schedule. About 84 percent of the 520 local areas, including those in Puerto Rico, completed their person interviewing at least 2 weeks before the scheduled end of the operation in their areas. The faster progress was partly due to the Bureau's ability to conduct about 28 percent of its national workload by telephone, compared to the 13 percent that it expected. Telephone interviews are faster and cheaper than in-person visits. But not all areas completed interviewing nearly as quickly; nine areas completed over one-fifth of their interview workloads during the final 2 weeks of person interviewing in their areas.

The quality assurance program for person interviewing appears to have generally covered the interview caseload as required, and, based on the results of the program, the Bureau assumes that about one-tenth of 1 percent of all cases nationally would have failed the program, having been believed to have been falsified. According to Bureau data, about another 2 percent of cases nationally would have failed quality assurance had the Bureau treated interview errors from other sources, such as honest mistakes, similarly to those attributed to falsification. The quality assurance program consisted of interviewers telephoning or visiting households of completed cases to verify whether the initial interview had actually taken place and reinterviewing in cases where falsification was suspected. Under the Bureau's quality assurance guidelines, 5 percent of the person interviewing cases were to be selected randomly for review. Another 5 percent of the caseload was to be judgmentally selected by quality assurance supervisors looking at a variety of production indicators. Other operational measures of data quality included the A.C.E. response rate as well as the rate at which data were collected from proxy respondents. The Bureau achieved an A.C.E. interview response rate of almost 100 percent of the occupied inhabitable housing units, which helped ensure a more accurate measurement of census errors. However, about 5 percent of cases were completed with proxy respondents, such as neighbors, and about 5 percent of cases received incomplete data during their interviews. The Bureau has found that the data collected from proxy respondents are generally less reliable than data collected from members of the respondents' households. The proxy rate varied locally; in 49 areas over one-tenth of the work was completed with proxy interviews. Local

variations in these measures could affect the accuracy of A.C.E. estimates of census undercounts. It will be important for the Bureau to conduct additional research to determine the relationship, if any, between these operational measures and the accuracy of A.C.E. estimates of census undercounts.

Early in the person interviewing operation, the Bureau dealt with an unexpected problem with its automated work management system, which allows supervisors to selectively reassign work among interviewers. According to Bureau officials, a programming error resulted in the unintended duplication of some cases being reassigned during the first 2 weeks of the field operation. However, according to Bureau A.C.E. officials, the Bureau addressed the underlying programming error within 2 weeks, and the operation proceeded on schedule. Bureau A.C.E. officials reported to us that regional supervisors later deleted the duplicate cases.

The address list used during person interviewing had fewer nonexistent listings than did the lists used by the major census questionnaire delivery operations. An accurate address list is important because it prevents unnecessary and costly efforts to locate nonexistent residences. The A.C.E. operations preceding the A.C.E. field survey were designed to flag addresses on the A.C.E. list that had been deemed questionable when compared to the initial census address list and later verified by field operations as nonexistent. As a result, during person interviewing, 1.4 percent of housing units to be interviewed were found not to exist. By comparison, about 9 percent of the census forms were returned as undeliverable as addressed during the nationwide mail-out of census questionnaires.

Although the Bureau implemented controls to keep the nonresponse follow-up operation separate from person interviewing, the assumed independence of the census and A.C.E. was put at risk because another follow-up operation, intended to improve census coverage, overlapped with person interviewing. We interviewed eight regional directors or members of their staffs who reported that they were not implementing controls to keep the other follow-up operation from overlapping with person interviewing, similar to those implemented to keep nonresponse follow-up separate from person interviewing, primarily because they viewed additional communications as possibly contributing to the risk of compromising independence. The Bureau told us that there never was an intent or requirement to implement such operational controls. Each of the regional directors we spoke with said that no significant overlap was taking

place in the field. Headquarters officials, based on their assessment of related studies of a similar overlap in 1990, decided in 1999 that the overlap would not adversely affect data quality. Yet information is not available on the extent to which interviews overlapped or on whether the independence assumption was operationally satisfied. Thus it will be important for the Bureau to do additional research on the extent of the overlap.

Overall, the Bureau overcame a number of operational challenges that could have undermined the success of A.C.E. person interviewing. Still, local variations in operational measures could have had an effect on A.C.E. calculations, Bureau definitions of quality assurance “failure” excluded most of the interviewing errors detected by the quality assurance programs, and gaps existed in the controls to prevent overlap between person interviewing and certain census follow-up operations. Thus, we recommend that the Secretary of Commerce ensure that the Bureau’s ongoing 2000 Census and A.C.E. evaluation efforts examine (1) whether any operational measures have a significant relationship with the accuracy of A.C.E. estimates of census undercounts, (2) broadening measures of quality assurance failure to include errors from sources other than falsification, and (3) issues related to controls, such as the extent of any overlap between census follow-up operations and A.C.E. person interviewing and whether additional controls to minimize overlap between census follow-up and person interviewing in the future can help ensure their independence.

In commenting on a draft of this report the Bureau provided minor technical corrections and additional information and also clarified some of our key points and recommendations, which we have reflected in this final report and comment on in more detail in appendix I where the Bureau’s comments are reprinted.

Scope and Methodology

To meet our objectives and review the implementation of person interviewing, we examined relevant Bureau program and research documents, such as procedures memorandums and analysis of Census tests. Further, we reviewed data from the Bureau’s “cost and progress” management information system, which Bureau officials used to monitor the conduct of census and A.C.E. operations. To help validate and expand on the cost and progress data, we interviewed Bureau headquarters and regional officials. We also interviewed key Bureau officials from headquarters and, where applicable, regional officials responsible for the planning and implementation of the person interviewing operation. Although

we verified with Bureau officials that the data were final, we did not independently verify data contained in the Bureau's cost and progress management information system.

To obtain a local perspective on how person interviewing was implemented, we interviewed temporary A.C.E. workers in 12 locations, covering over 60 local census office areas (out of the 520 in the United States and Puerto Rico) and corresponding to 8 of the 12 census regions (Atlanta, Boston, Charlotte, Chicago, Dallas, Denver, Los Angeles, and Seattle).² To provide further context, we also interviewed A.C.E. managers of seven of these regions. We selected these areas primarily for their geographic dispersion, variation in type of enumeration area, and their proximity to our field offices. The results of the visits could not be generalized to all person interviewing.

In addition to these field locations, we performed our audit work on eight of the Bureau's A.C.E. regions at Bureau headquarters in Suitland, MD; as well as in Washington, DC, from June 2000 through January 2001, in accordance with generally accepted government auditing standards. On September 7, we requested comments on a draft of this report from the Secretary of Commerce. On October 5, 2001, the Secretary of Commerce forwarded written comments from the Bureau (see appendix I), which we address in the "Agency Comments and Our Evaluation" section of this report.

Person Interview Operation Generally Completed on Schedule

The Bureau appears to have generally completed person interviewing according to its operational schedule. Failure to collect data in a timely manner could have reduced the interview completion rate or increased the Bureau's dependence on less reliable sources of data, such as proxy data, thus reducing the quality of data collected. In addition, the Bureau believes that the more time that passes from Census Day (April 1) to the time of the survey interview, generally the more likely that the survey respondent will err in his or her recall of Census Day information. Finally, data processing and other operations depended on the data from person interviewing, and

²The Census Bureau has regional offices in Atlanta, GA.; Boston, MA; Charlotte, SC; Chicago, IL; Dallas, TX; Detroit, MI; Denver, CO; Los Angeles, CA; New York, NY; Philadelphia, PA; Kansas City, KS; and Seattle, WA. The nine local census offices in Puerto Rico are administratively reported as in the Boston region.

could have been delayed had person interviewing not been completed on time.

Success With Telephone Interviewing Decreased Door-to-Door Workload

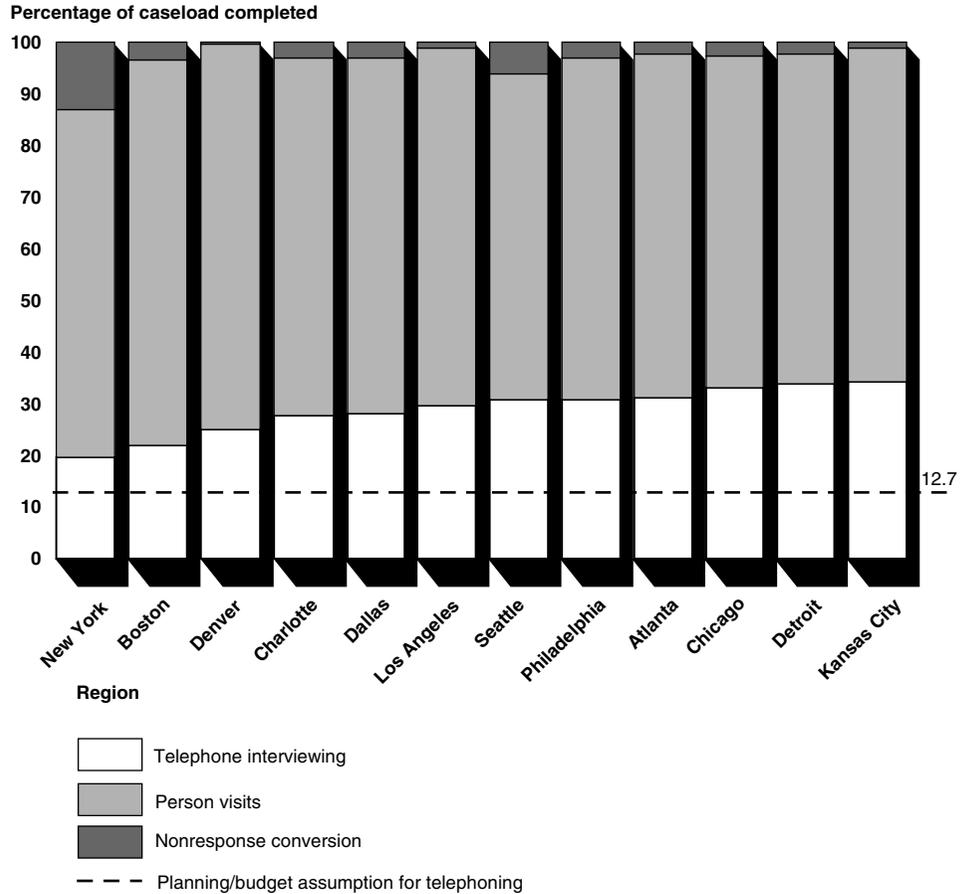
About 84 percent—439—of 520 local census office areas completed all of their fieldwork at least 2 weeks before the end of the operation in their areas. By the deadline for completing all person interviewing, September 1, 2000, only 45 cases (out of over 314,000 nationally) remained to be completed, and they were all in the area of a single local census office, which had received an extension of its deadline.

The timely completion of person interviewing was due in part to the Bureau's ability to conduct a much higher share of the person interviewing caseload by telephone than it had anticipated. Although the Bureau anticipated that about 40,000 cases (about 13 percent) of the person interviewing caseload would be completed by telephone, the actual amount was much higher—about 90,000 cases (over 28 percent). Bureau officials informed us that more people provided their telephone numbers on their census returns and more people returned their census forms than the Bureau had anticipated. The telephone phase of the operation was limited to cases in which households had provided telephone numbers with their census responses—about 40 percent of the roughly 314,000 total person interviewing cases.

As figure 2 illustrates, the share of the workload completed by telephone varied across regions, ranging from 19 to 34 percent. It also varied across local census office areas, ranging from less than 1 to over 55 percent. Bureau officials explained that this variation was related to the eligibility criteria, which further limited telephoning to households with city-style addresses (for example, 123 Main Street) that were not in small multiunit dwellings.³ Completing interviews over the telephone reduces the travel time for interviewers and can thus decrease the cost of each interview completed.

³Housing units in city-style areas receive their mail addressed predominantly to a building number and street name; housing units in noncity areas receive mail delivered primarily to other styles of addresses, including post office boxes and rural route addresses.

Figure 2: Proportion of Person Interviewing Workload Completed by Telephone Interviews by Census Region



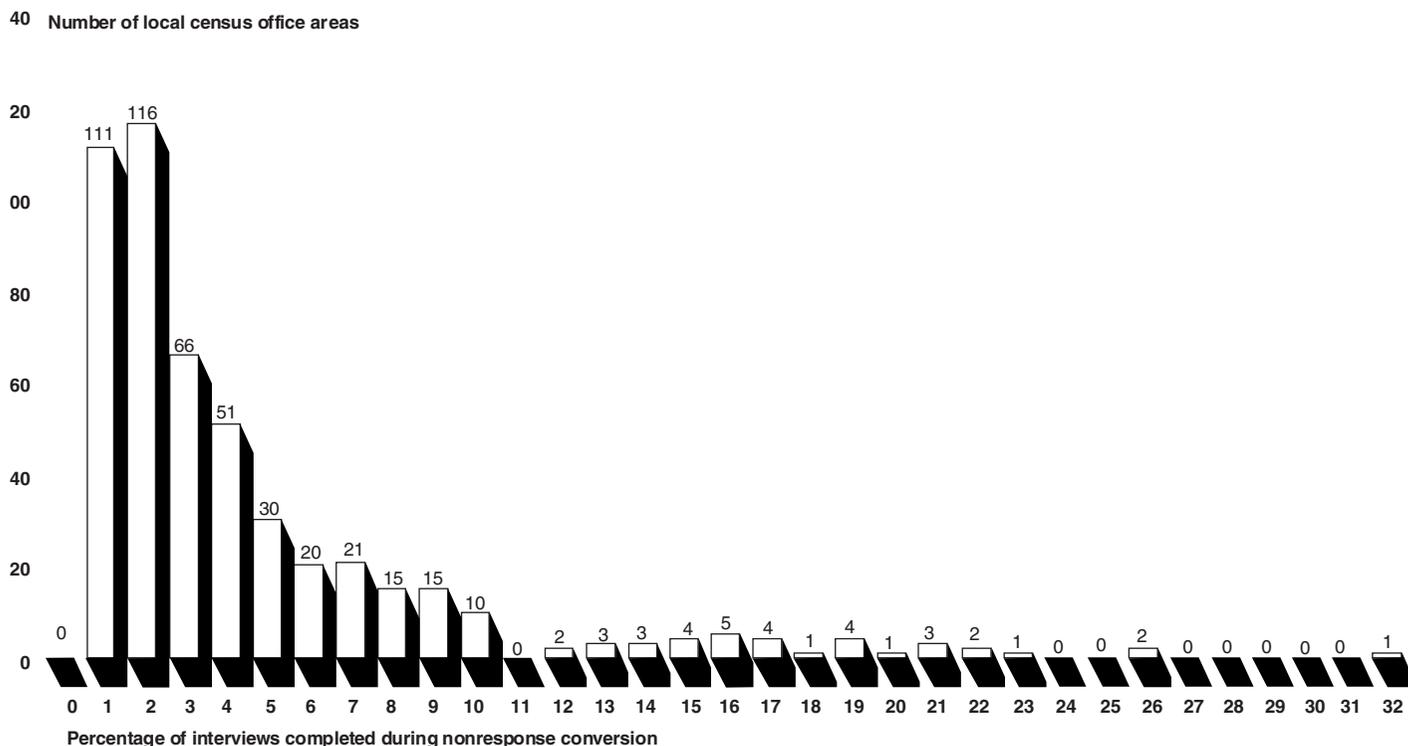
Source: GAO analysis of U.S. Census Bureau data.

Some Areas Completed 10 Percent of Their Workload During the Last 2 Weeks of the Operation

Person interviewing did not progress equally in all local census office areas. About 3 percent of the national person interviewing caseload had to be reassigned to nonresponse conversion, which took place in the final 2 weeks of interviewing in a given area. This compares to the about 2 percent of the person interviewing that was completed during an ad hoc nonresponse conversion operation during 1990. As shown in figure 3, about 36 areas out of the 520 had over 10 percent of their cases reassigned to this last phase, and 9 areas had more than one-fifth of their caseloads reassigned to this phase. The New York region had almost 13 percent of its

caseload referred to nonresponse conversion. Bureau officials told us that, in response, the New York region brought in professional interviewers from other nondecennial survey work to conduct its nonresponse conversion, hoping to ensure a high-quality interview process.

Figure 3: Distribution of 520 Local Census Office Areas' Share of Person Interviewing Completed During Nonresponse Conversion



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

Quality Assurance of Person Interviewing Appeared to Progress on Track

During the person interviewing operation, the Bureau carried out a quality assurance program, which focused primarily on detecting interviews falsified by interviewers. Bureau officials designed the person interview quality assurance program to detect when interview results had been submitted but the interview had not been done, because Bureau research suggested that the most common type of falsification was the falsification of an entire interview. Outside the quality assurance program, a number of

operational indicators associated with data quality in the past suggest that data quality may have varied locally.

Quality Assurance Workload Guidelines Satisfied

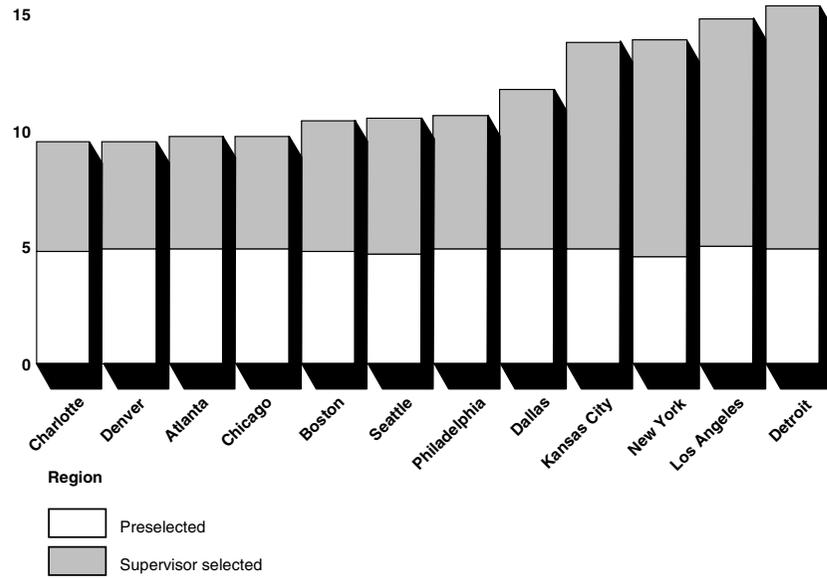
Under the Bureau's quality assurance program, regional offices were to telephone or visit a 5-percent random sample of all person interview cases to determine whether an initial person interview had actually taken place. Further, according to headquarters quality assurance managers, regional quality assurance managers were to select about another 5 percent relying on automated "outlier reports" and other criteria. For example, supervisors were required to select additional cases for quality review when outlier reports showed that an interviewer had a relatively high percentage of vacant housing units or interviews conducted at unusual hours and thus might be falsifying data. Every interviewer was to have at least one case covered by quality assurance. As an additional check, the Bureau provided quality assurance supervisors with reports on respondents' names so that they could look for indicators of possible falsification by interviewers, such as names of famous characters/people or multiple respondents with the same name.

Figure 4 illustrates the cumulative share of each of the 12 regions' person interviewing caseload that was reviewed by the quality assurance program. By the end of the operation, the national share exceeded 11 percent, and each of the 12 regions was near or exceeded the target ratio of about 10 percent of the person interview workload. Headquarters A.C.E. quality assurance managers said that the percentage of the interview caseload selected for quality assurance review was expected to vary depending on a number of local circumstances. For example, where quality reviews raised the suspicion of fraudulent interviews, supervisors were to select more of those interviewers' cases for review, further increasing the percentages reviewed by quality assurance in those areas. As figure 4 illustrates, in some of the regions with higher percentages of cases suspected of falsification by interviewers, supervisors did indeed select a higher percentage of cases for review. Nationally, less than 3 percent of the quality assurance cases were suspected of falsification, although across regions the percentage varied from about 1.3 to 5 percent. In comparison, Bureau evaluations of a 1998 dress rehearsal of person interviewing reported suspected falsifications of from about 0.9 to 3.1 percent of the quality assurance cases at the three different rehearsal sites. Although the sites of the dress rehearsal were not representative of the whole nation, they provide a reasonable benchmark for the 2000 census. As the Bureau noted in its response to our draft report, none of the sites were exceptionally

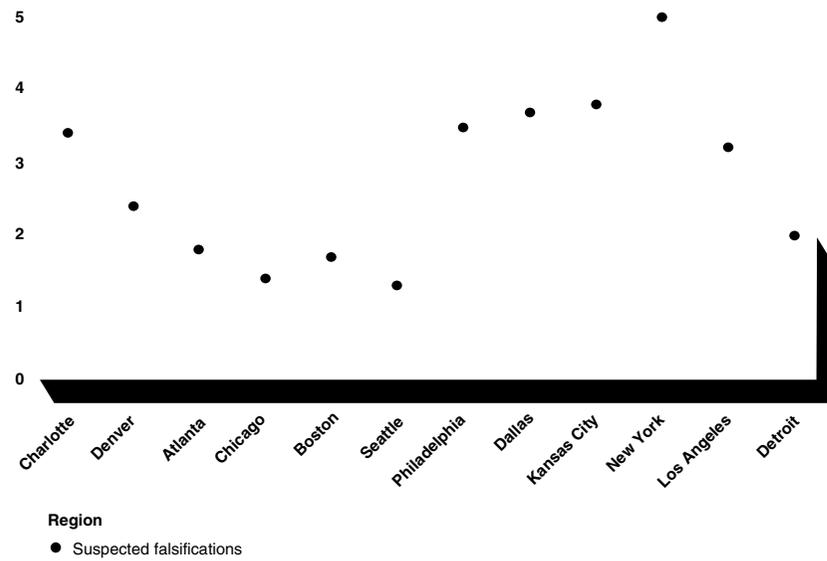
hard-to-enumerate areas, which tend to have higher rates of falsification. So, that 6 of the 12 regions, each with hard-to-enumerate areas, had their rates fall into this range, demonstrates in part that the Bureau's person interviewing experienced low rates of suspected falsification in 2000.

Figure 4: Person Interviewing Quality Assurance by Census Region

20 Percentage of person interviewing workload



6 Percentage of quality assurance workload



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

Each case suspected of falsification was to be reinterviewed, as was the entire workload of any interviewer found to have actually falsified a case. A total of 1,004 cases (2.8 percent of the quality assurance caseload) had their data replaced by these reinterviews. After further investigating the cases suspected of falsification, Bureau officials believed that about 0.1 percent of the randomly selected quality assurance caseload stateside (0.2 percent including Puerto Rico) was falsified and assumed that this percentage is generalizable to the entire A.C.E. sample. The percentage of the replaced interviews that contained errors due to honest interviewer mistakes, poor respondent recall, or reasons other than falsification was not reported by the Bureau's quality assurance program as part of the failure rate. But data that the Bureau provided later show that 2.1 percent of all randomly selected cases stateside (and 2.1 percent including Puerto Rico) were replaced. We discussed the utility of the Bureau defining, measuring, and reporting a broader measure of quality assurance failure—including failure for reasons other than falsification—with the Associate Director for Decennial, and he concurred that the Bureau should consider this in the future.

All cases in the final phase of person interviewing, nonresponse conversion, were excluded from the quality assurance program. Headquarters officials said that because (1) the quality assurance was intended primarily to identify falsified interviews, (2) only the most experienced workers were used in the final phase of interviewing, and (3) the most experienced workers falsify less, no quality assurance was deemed necessary for that phase. They also pointed out that there would not have been time to check the cases completed on the last days of the operation, and a relatively small percentage of the total person interview caseload fell into this phase of the operation. According to Bureau data, the nonresponse conversion phase had a workload of about 10,000 cases, or about 3 percent of person interviewing cases nationwide.

Indicators of Quality Varied Locally

According to Bureau data, person interviewing collected information on current residents in almost 100 percent of the cases for housing units that existed, were inhabitable, and were not vacant. Following the change in the census design after the 1999 Supreme Court ruling, the Bureau no longer specified a goal or target for person interviewing response rate.⁴ However,

⁴*Department of Commerce v. U.S. House of Representatives*, 525 U.S. 316 (1999).

this response rate exceeded the 95 percent expectation expressed by the Bureau's Field Directorate in internal memorandums in 1997, as well as the 98 percent target the Bureau set prior to the Supreme Court ruling and the response rate of 98.6 percent during a similar survey during 1990—the 1990 Post-Enumeration Survey. All regions exceeded 99.7 percent, and only three local census office areas had person interview response rates lower than 98 percent.

However, not all of these interviews obtained complete information on the household. About 5.1 percent of cases nationally were recorded as “partial” interviews—interviews missing either the age, sex, or Census Day residency status for one or more household members.⁵ The regions varied in their rate of partial interviews from about 1 percent to over 8 percent. When interviews were not complete, the missing data were to be provided through statistical methods, and this would be a source of error in the resulting A.C.E. calculations. And, as we have reported previously, low interview completion rates could have resulted in some segments of the population being underrepresented in the A.C.E. data, adversely affecting the accuracy of any A.C.E.-based adjustments.⁶ Bureau officials believe that numerous studies over the years have shown that their procedures for dealing with missing data have acceptable error levels.

Furthermore, Bureau data show that about 5 percent of the household interviews were completed with proxy respondents, such as neighbors. According to Bureau research, proxy interviews do not generally provide information as reliable as interviews with household members, and this can be a source of error in A.C.E. calculations. Proxy interviews are also more likely to provide only partial information. The 2000 proxy rate exceeded the 2-to-3-percent proxy rate experienced during the 1990 Post-Enumeration Survey. The Boston region reported as little as 2 percent of its caseload completed by proxy, and the New York region had over 8 percent completed by proxy.

Local variations in data quality may affect the accuracy of A.C.E. results for some segments of the population. Although national level data are important for determining broad trends, they often mask implementation

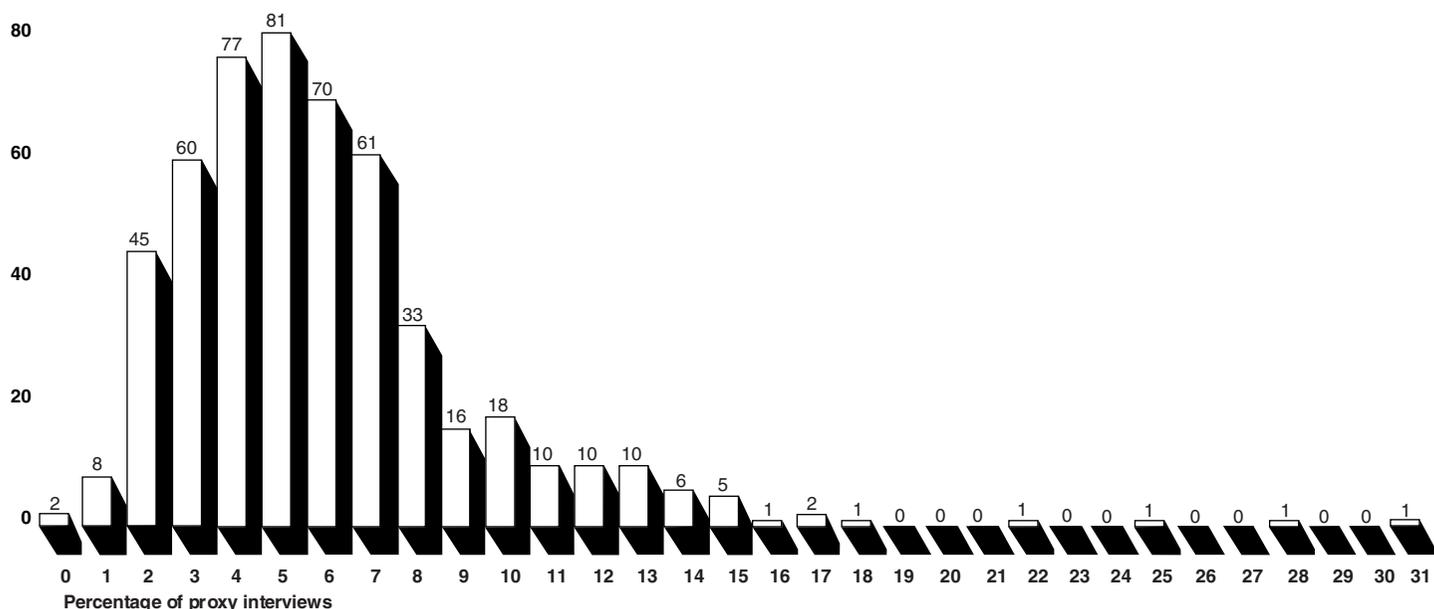
⁵According to the Bureau, the rate of partial interviews was calculated as the number of partial interviews divided by the sum of the partials, completes, and noninterviews.

⁶GAO/GGD-98-74 *2000 Census: Preparations for Dress Rehearsal Leave Many Unanswered Questions*, March 26, 1998.

challenges occurring at the local level. For example, as figure 5 shows, most local census office areas relied less on proxies than the national effort did, but 49 had to complete over 10 percent of their total caseloads with proxy respondents. Bureau officials said that the local areas with the highest proxy rates tended to be dense urban areas, such as in New York City, where buildings may have had restricted access, and interviewers had to rely on apartment managers for information.

Figure 5: Distribution of Proxy Rates for 520 Local Census Office Areas

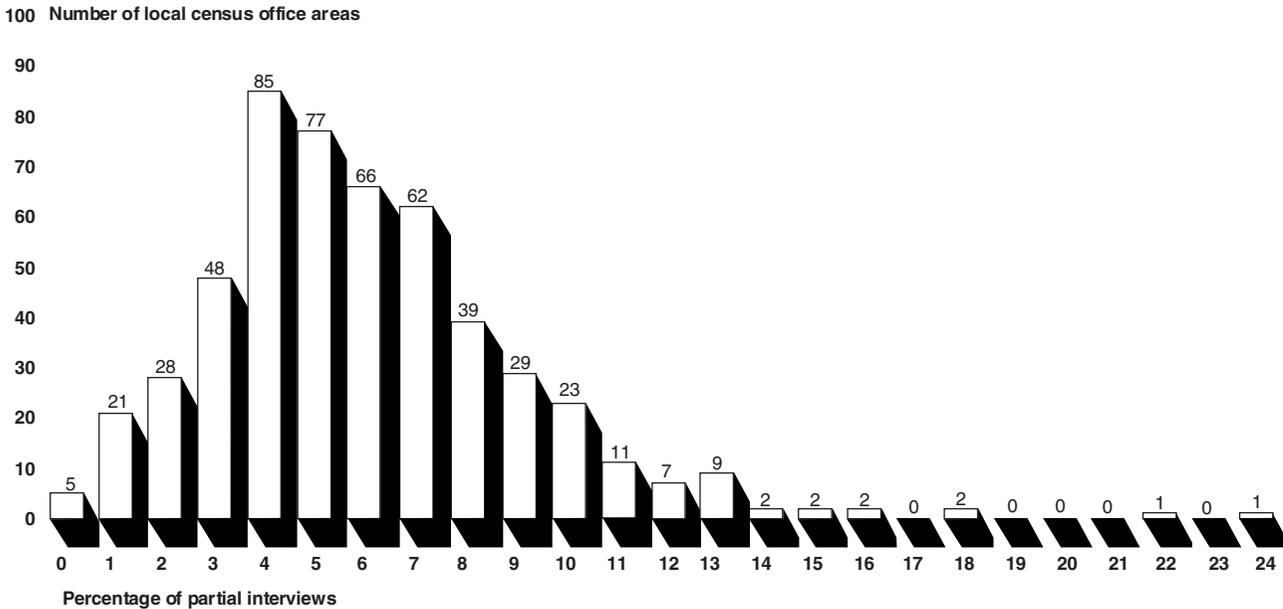
100 Number of local census office areas



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

Similarly, most local census office areas had rates of partial interviews near or below the national rate of about 5.1 percent, but 37 had rates exceeding 10 percent, as shown in figure 6. Bureau officials explained that many of these areas with the highest partial interview rates were areas with higher proxy rates as well because proxy interviews are less likely to provide complete data. Most local areas were near or below the nation's 0.1 percent final nonresponse rate on person interviewing cases, although one area had a nonresponse rate of 2.3 percent, and nine local areas exceeded 1 percent.

Figure 6: Distribution of 520 Local Census Office Partial Interview Rates



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

Most local census office areas completed less than or about the nationwide 3 percent of their caseload during nonresponse conversion. As shown in figure 3, the percentage of the person interviewing workload completed during nonresponse conversion exceeded 10 percent in 36 of the 520 local census office areas and 20 percent for 9 areas. As discussed earlier, nonresponse conversion was not subject to the quality assurance program, although the Bureau relied on its best workers for this stage of interviewing. The Bureau reports that quality assurance was not done during nonresponse conversion because that stage involved getting cooperation from uncooperative respondents, and the later A.C.E. field operation, person follow-up, would serve as a form of quality assurance on these interviews.

Programming Errors Temporarily Hindered Management of Person Interviewing

Early in the person interviewing operation, the Bureau experienced and resolved problems with a critical function in its automated work management system that was to allow supervisors to selectively reassign work among interviewers. The software was to enable supervisors to reassign cases that had either been sent to them for review or that needed

to be reassigned from a laptop computer that was either broken or had been issued to someone who was no longer interviewing. For cases being reassigned that had not been flagged for supervisory review, the software was to ask the supervisor whether to disable the cases on the original laptop. If the cases were being reassigned to a different interviewer, supervisors were to disable the cases.

However, according to Bureau officials, the software contained errors in two different places. One error resulted in cases not being disabled from an interviewer's laptop even after the supervisor attempted to disable them. Another error resulted in certain cases being reassigned automatically, again without them being disabled on the original laptop. Both problems resulted in duplicate records on the laptops, which required supervisors to individually review and delete cases. According to Bureau officials, this confused some temporary staff and their supervisors and created some work inefficiencies. For example, some households received unplanned multiple visits by different interviewers. However, according to Bureau A.C.E. officials, the Bureau addressed the underlying programming error within 2 weeks, and the operation proceeded without a reported recurrence of this problem.

Person Interviewing Address List Quality Appeared to Be Better Than the Census

An accurate address list avoids unnecessary and costly efforts to locate nonexistent residences. A measure of address list quality is the percentage of addresses that are nonexistent or "undeliverable" because interviewers were unable to locate housing units at the listed addresses. During person interviewing, 1.4 percent of housing units to be interviewed were deemed to not exist, and this was less than for other major census questionnaire delivery operations.

Table 1 illustrates that the two primary census questionnaire delivery operations both experienced initial undeliverability rates greater than the share of nonexistent housing units during person interviewing. In comparison, person interviewing during the 1990 Post-Enumeration Survey also encountered a higher share of its caseload being undeliverable than did person interviewing in 2000.

Table 1: Deliverability of Initial Address List for Major Census Operations

Operation	Approximate number of addresses on respective initial Bureau lists	Percentage undeliverable as addressed^a
Update/Leave Census 2000	22,000,000	5.2 %
Mail-out Census 2000	99,000,000	9.1% ^b
Person Interviewing, A.C.E. 2000	314,000	1.4%
Person Interviewing, Post-Enumeration Survey 1990	170,000	2.6%

^aIncludes addresses to which questionnaires could not be hand delivered because the housing units were found uninhabitable, the addresses did not exist, or the structures did not contain housing units as well as addresses that the U.S. Postal Service returned to the Bureau as undeliverable.

^bExcludes about 1.6 million questionnaires returned by the U.S. Postal Service as “undeliverable as addressed” due, for example, to incorrect zip codes or lack of residential delivery in the area, but that were successfully redelivered by census workers.

Source: U.S. Census Bureau.

The list of addresses visited during person interviewing benefited from earlier A.C.E. operations that (1) independently canvassed all addresses in A.C.E. areas, (2) compared the initial A.C.E. address list to the initial census address list, (3) reconciled any differences by field visits, (4) flagged nonexistent addresses in A.C.E. sample areas, and (5) entirely relisted some areas that were canvassed improperly. Person interviewing attempted to visit only addresses that had been confirmed to exist in A.C.E. sample areas.

Census Follow-up Operations Overlapped With Person Interviewing in the Field

As noted earlier, the Bureau’s procedures called for it to go to great lengths to ensure that A.C.E. operations were kept independent of the census operations to avoid biasing A.C.E. estimates. For person interviewing, this meant conducting the operation after the Bureau completed nonresponse follow-up activities in a local census office area; implementing controls to prevent their overlap; sharing status information about nonresponse follow-up with A.C.E. managers; and managing field activities out of 12 separate regional census offices, independent of the 12 regional census centers managing the rest of the census.

However, in response to the 1999 Supreme Court ruling against the planned use of sampling to generate population data for reapportioning the House of Representatives, the Bureau reintroduced a census follow-up operation

intended to improve census coverage in part by sending enumerators to households that were added to the census address list late and thus may have been missed by earlier census operations. The schedule of this operation, known as “coverage improvement follow-up,” overlapped the beginning of person interviewing, thus increasing the risk that it would violate the independence assumption. According to Bureau officials, a similar operation had overlapped person interviewing for the Post-Enumeration Survey in 1990, and delaying person interviewing further from Census Day would have increased the risk that respondents would not reliably recall their Census Day data.

The risk of violating the independence assumption was increased further when the workload of the follow-up operation increased over what was projected. In December 1999 the estimated workload volume for coverage improvement follow-up was about 7.7 million addresses. Most of these were to have verified the vacancy or nonexistence of housing units previously marked vacant or for deletion. But up to about 0.8 million addresses were more likely to have enumeration interviews take place. A.C.E. designers believed that the number of coverage improvement cases in a given area would not be enough to affect A.C.E. data collection. On the basis of the actual number of addresses being covered by the operation, by June 2000 this number had risen to about 2.3 million addresses. In addition, after person interviewing had begun, the Bureau decided to revisit every census household for which the population count was unknown. According to Bureau sources, there were about 0.7 million such households.

Although the Bureau had strict controls to prevent person interviewing from going door-to-door in areas where census nonresponse follow-up—the primary census field follow-up operation—was still under way, the controls did not apply to other census follow-up operations, such as coverage improvement follow-up. Automated work management rules were to prohibit person interviewing field visits from beginning in a local census office area prior to the earlier of either (1) 100 percent completion of nonresponse follow-up in that local census office area or (2) 1 week after 90 percent completion of nonresponse follow-up in all A.C.E. clusters in that local census office area. In addition, A.C.E. management had access to “early warning reports” that provided the daily status of nonresponse follow-up in each area. According to the Bureau, exceptions to the start rules had to be approved in headquarters, and the only software that would allow an earlier start to personal visits was located at headquarters. The Bureau also informed us that the regional offices did not have the ability or the authority for exceptions to be implemented, as any changes required at

least Assistant Division Chief level approval. According to the Assistant Field Division Chief for Evaluation and Research, to his knowledge, no such approvals had been given. He said that these rules did not apply to other follow-up operations.

Each of the eight regional directors (there are 12 in all) or their deputies with whom we spoke regarding A.C.E. independence said that no significant overlap occurred between concurrent census follow-up and person interviewing. However, most of them believed that some overlap was likely, and none of them could be certain of the extent of any actual overlap. Moreover, all of these regional directors or their A.C.E. management staffs also reported not having any communication from the census side of their operations to the A.C.E. operations on the status of these follow-up operations beyond that on the nonresponse follow-up work, underscoring their inability to control the possible overlap of census and A.C.E. fieldwork.

Prior Bureau research on sample data collection detected few possible effects of overlap; and, at the time the coverage improvement follow-up operation was reintroduced, Bureau officials concluded that there was no significant risk to independence. The Chief of the Bureau's Decennial Statistical Studies Division said that on the basis of his experience and understanding of prior Bureau research, the small risk of compromising independence was worth taking to reduce the risk of increased errors from delaying person interviewing until the Bureau completed coverage improvement follow-up. He and other headquarters officials we interviewed were unaware of any Bureau attempts to determine the extent of any possible interview overlap in 2000, which might demonstrate whether A.C.E. assumptions were operationally supported. The Bureau recently completed, as part of its Census 2000 evaluation program, a study intended to detect significant differences between the census responses in comparable A.C.E. and non-A.C.E. blocks. The study found no differences it deemed significant.

Conclusions

The Bureau largely overcame significant challenges that could have undermined the person interviewing operation. Notably, the Bureau completed the person interviewing data collection on schedule and in accordance with its general guidelines for quality assurance coverage. The Bureau also demonstrated its ability to overcome the limited technical challenges it confronted. Furthermore, the series of A.C.E. address operations, as designed, appeared to effectively remove nonexistent

housing units and addresses from the person interviewing caseload, thus reducing an otherwise inefficient use of interviewing time and resources.

Still, the Bureau's experience in implementing person interviewing highlights areas where additional research might lead to improvements if the Bureau conducts a similar operation for the 2010 Census. For example, certain operational challenges may have contributed error to final A.C.E. estimates of census undercounts and overcounts. The Bureau experienced variation at the local level in how person interviewing was carried out, in terms of response rates, proxy rates, and partial interview rates. As we have reported before, if the local census office areas with the worst values of each of these measures have populations that are typically hard to count in the census, these segments of the population may be underrepresented in the A.C.E. data, possibly leading to inaccurate reflections of these population segments in A.C.E.-based adjustments. The Bureau plans to evaluate the relationship between operational measures, such as proxy rates, and how well A.C.E. data match to census data. The results of these evaluations, and others, will provide an important basis for planning an improved 2010 census and evaluation survey.

Further, although Bureau data show that the person interviewing quality assurance program met its objectives, the program focused primarily on identifying falsification and reported failure rates based solely on cases believed to have been falsified. As the Bureau looks to improve its interviewing experience further, a broader definition of quality assurance failure to include interviews the Bureau reinterviewed and replaced for other reasons would provide a more complete measure of interviewing quality.

Finally, the same controls and sharing of status information to ensure independence between the census nonresponse follow-up operation and A.C.E. person interviewing were not applied or did not take place with other census follow-up operations, thus increasing the risk of compromising the independence of A.C.E. A relatively small part of the census follow-up workload was not subject to control over its possible overlap with A.C.E. person interviewing and thus the magnitude of this influence may have been small nationally; however, it could potentially have been significant in some local areas. To that extent, the A.C.E. assumptions may not apply equally in all areas or for all segments of the population, with possible adverse effects on the accuracy of A.C.E. calculations.

Since the Bureau will likely use an evaluation survey in 2010, perhaps similar to A.C.E., it will be important for the Bureau to learn the lessons from the 2000 Census that can be incorporated into the planning for 2010.

Recommendations for Executive Action

As the Bureau documents its lessons learned from the 2000 Decennial Census and as part of its planning efforts for 2010, we recommend that the Secretary of Commerce conduct research that

- determines the relationship, if any, between operational measures of person interviewing, such as proxy rates, and the accuracy of A.C.E. estimates of census undercounts as planned;
- determines how best to define, measure, and report interview quality failure rates that include interviews rejected for all reasons, and not just for a subset of reasons such as falsification;
- determines and documents the extent, if any, of the actual overlap between census follow-up operations and A.C.E. person interviewing in 2000;
- determines whether sufficient overlap may have occurred to violate the independence assumption;
- determines whether increasing the flow of status data on specific decennial follow-up operations to the managers of independent surveys can help ensure the independence of such surveys, particularly when such operations are scheduled to overlap in the field; and
- determines what additional steps or controls to preserve the independence of census follow-up and person interviewing, if any, could be implemented for other census follow-up operations that collect enumeration data and are scheduled contemporaneously with person interviewing.

Agency Comments and Our Evaluation

The Secretary of Commerce forwarded written comments from the Bureau on a draft of this report. (See appendix I). The Bureau provided minor technical corrections and additional information. The Bureau also offered clarification on some of our key points and recommendations, which we have reflected in this final report and comment on in more detail in appendix I.

Regarding our finding that census follow-up operations overlapped with person interviewing in the field, the Bureau provided additional information on its decision to permit overlap between census coverage

improvement follow-up and the A.C.E. person interviewing operations. We recognized this context, and revised the draft to better reflect it. Nevertheless, while the Bureau's response justifies its not adding any controls or communications or changing any procedures when it noticed that the increase in the follow-up workload was over what was projected, as we note in the report, the increase in workload increased the risk that the independence assumption was violated. There may still be opportunities to implement steps in the future to help ensure the independence of such surveys.

The Bureau commented that our conclusion linking variations in data quality to possible effects on the accuracy of A.C.E. results was unsubstantiated and suggested wording it as a question. In our draft report, we had raised the link as a possibility and then recommended that the relationship, if any, be determined between operational measures and the accuracy of A.C.E. estimates. We believe that this conclusion is logical given other Bureau reporting linking data quality measures such as missing data rates to possible errors in A.C.E. results.⁷ We have also reported on this issue in the past.⁸

The Bureau said that our conclusion that the influence of census and A.C.E. overlap may have been significant in some local areas was unsubstantiated. We were unable to conclude whether significant overlap had occurred or not. As we noted in our report, the Bureau officials we interviewed were unaware of any Bureau attempts to determine the extent of any possible interview overlap in 2000, which might have demonstrated whether A.C.E. assumptions were operationally supported in the field. Without such evidence regarding the extent of the overlap, and given the anecdotal evidence, which the Bureau cites in its response and we mentioned in the draft report, that some overlap did occur, we view the conclusion that the overlap may have been significant in some areas as appropriate. We revised the text to more clearly state, however, that the effect of the overlap is a potential one.

⁷Bureau of the Census *Report of the Executive Steering Committee for Accuracy and Coverage Evaluation Policy on Adjustment for Non-Redistricting Uses* (October 17, 2001).

⁸GAO/GGD-98-74 *2000 Census: Preparations for Dress Rehearsal Leave Many Unanswered Questions*, March 26, 1998 and GAO/GGD-97-142 *2000 Census: Progress Made on Design, but Risks Remain*, July 14, 1997.

In responding to our recommendations (1) to determine the relationships, if any, between operational measures and the accuracy of A.C.E. estimates, as planned, and (2) to determine and document the extent of overlap between census and A.C.E. in 2000, the Bureau acknowledged the importance of extensive evaluations of A.C.E., and referred to the evaluation it is undertaking. We look forward to reviewing this evaluation when it is complete.

Since receiving comments from the Bureau, we added one additional recommendation. The basis for this new recommendation centered on our finding that the Bureau's quality assurance program did not report fully on the percentage of the interview workload replaced by the quality assurance interviews. Although we recognize that the quality assurance program was designed primarily to detect falsification, the definition of quality assurance "failure" used by the Bureau excluded the sources of error other than falsification. After receiving the Bureau's response, we discussed this with the Associate Director for Decennial at the Bureau, who concurred that the Bureau should consider a broader definition in the future. We have added an additional recommendation for executive action accordingly.

In responding to our recommendation to determine whether sufficient overlap occurred to violate the independence assumptions, the Bureau referred to its recent evaluation of the possible contamination of census data collected in A.C.E. blocks, as well as several other similar studies throughout the decade. Some of these studies find weak or only limited indications of contamination of census data in prior censuses, and they all conclude that there was no systemic contamination of census data. The Bureau's most recent evaluation, which is consistent with our recommendation, was released after our audit work was completed. We have revised the draft accordingly.

We are sending copies of this letter to other interested congressional committees.

Please contact me on (202) 512-6806 if you have any questions. Other key contributors to this report are included in appendix II.

A handwritten signature in black ink, reading "J. Christopher Mihm". The signature is written in a cursive style with a large, stylized "J" and "M".

J. Christopher Mihm
Director
Strategic Issues

Comments from the Secretary of Commerce

Note: GAO comments supplementing those in the report text appear at the end of this appendix.



THE SECRETARY OF COMMERCE
Washington, D.C. 20230

OCT - 5 2001

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

Dear Mr. Mihm:

The Department of Commerce appreciates the opportunity to comment on the General Accounting Office draft report entitled *2000 Census: Coverage Evaluation Interviewing Overcame Challenges, But Further Research Needed*. The Department's comments on this report are enclosed.

Warm regards,

A handwritten signature in black ink, appearing to read "Donald L. Evans".

Donald L. Evans

Enclosure

Appendix I
Comments from the Secretary of Commerce

Enclosure

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

U.S. General Accounting Office (GAO) report entitled *2000 Census: Coverage Evaluation Interviewing Overcame Challenges, But Further Research Needed*

Comments on the Text of the Report

See comment 1.

It appears that figures cited throughout this report may be inconsistent. This could happen because GAO is using different sources for its data, or because the Census Bureau has data that are more recent than that available to GAO. Census Bureau officials will work directly with GAO to address any discrepancies of this nature.

See comment 2.
Now on p. 3.

Page 3, Figure 1— The “Housing unit matching” and “Person matching” boxes give no indication that both a large clerical operation and a large field operation are subsumed under that heading. We suggest that the boxes contain text such as:

Housing Unit Matching Clerical Match → Field Housing Unit Follow-up
--

Person Matching Clerical Match → Field Person Follow-up Interview
--

See comment 3.
Now on p. 5.

Page 5, paragraph 3— The word “randomly” should be added to the third sentence so that it reads: “Under the Bureau’s quality assurance guidelines, 5 percent of the persons interviewing cases were to be *randomly* selected for review.”

See comment 4.
Now on p. 6.

Page 6, paragraph 3— Prior to the start of Accuracy and Coverage Evaluation (A.C.E.) operations, the issue of the overlap of the census coverage improvement follow-up operation and A.C.E. Person Interview was discussed in a number of forums, including the A.C.E. Program Steering Committee. It was clear that to delay A.C.E. operations in a Local Census Office (LCO) until the census coverage improvement follow-up operation had concluded would introduce considerable risk. This includes recall bias due to increased time lag between “Census Day” and A.C.E. interviews and the ability to meet critical scheduled deadlines, including the

Appendix I
Comments from the Secretary of Commerce

production of final undercount estimates. A.C.E. Person Interview cases were not released for personal visit interviewing individually, or by cluster, but only by LCOs. It would have been impossible to manage the survey using a smaller unit of geography. It was a conscious decision, made long before interviewing began, to allow the overlap of the census coverage improvement operation and the A.C.E. Person Interview operation in the schedule.

The A.C.E. designers also believed that the number of coverage improvement cases in a given area would not be sufficient to have an effect on the A.C.E. data collection. There was never any intent or requirement to implement operational controls for independence between A.C.E. and census coverage improvement. The census coverage improvement follow-up operation work load was larger than predicted. The A.C.E. procedures, controls, and automation were not changed at the last minute to accommodate this unexpected increase. Census regional management correctly followed existing procedures. Ad hoc procedures, without headquarters' prior approval, were prohibited for A.C.E. In fact, this overlap was discussed in A.C.E. teleconferences, and the regions were told not to implement any ad hoc procedures to prevent the overlap. In the teleconferences, we did hear a few anecdotal accounts of an A.C.E. interviewer seeing a Coverage Improvement Follow-up (CIFU) interviewer in the field.

See comment 4.

See comment 5.
Now on p. 8.

Page 7, footnote— St. Louis, MO, should be deleted from the list of the Census Bureau's regional offices and replaced with Kansas City, Kansas.

See comment 3.
Now on p. 9.

Page 8, paragraph 1— It would be helpful to clarify this paragraph by pointing out that the only LCO in the Nation, excluding Puerto Rico, which had cases remaining after the September 1, 2000, deadline, was in Hiialeah, and this was due to an agreed-upon extension of the deadline for this LCO.

See comment 6.
Now on p. 9.

Page 8, paragraph 2— The report states that "Bureau officials told us that more people apparently provided their telephone numbers on their census returns than they had anticipated." We are not sure that this reason is the appropriate reason for the disparity. In addition to more people entering their phone numbers on the census forms, more people than expected mailed back their census form. Both factors very likely influenced the pool of available telephone numbers. Our planning figure for percent of the interviews to be completed by telephone was based on dress rehearsal results. Going into A.C.E. 2000, we realized there were many variables surrounding this percentage that were not known.

Appendix I
Comments from the Secretary of Commerce

See comment 7.
Now on p. 12.

Page 10, paragraph 2— The report states that “outlier reports showed that an interviewer had a relatively high percentage of vacant housing units or proxy interviews and thus might be falsifying data.” Proxy interviews are known indicators of insufficient data quality—not fabrication. Indicators of fabrication include missing telephone numbers, days with more than 13 cases, and so forth.

See comment 8.
Now on p. 12.

Page 11, last sentence—A comparison of suspected falsification rates between the dress rehearsal and 2000 is not really valid. The three dress rehearsal sites were not representative of the Nation. None of the dress rehearsal areas were exceptionally hard-to-enumerate areas. In household data collection, certain factors almost always go together. Hard-to-enumerate areas lead to higher nonresponse rates, more proxy interviews (in surveys that allow proxy responses), more missing data, and, unfortunately, higher rates of falsification.

See comment 9.
Now on p. 12.

Page 13, paragraph 1— The report states that “as figure 4 illustrates, in some, but not all, of the regions with higher percentages of cases suspected of falsification by interviewers, supervisors selected a higher percentage of cases for review. Every case suspected of falsification was to be entirely reinterviewed, as was the entire workload of any interviewer found to have actually falsified a case.” This is very misleading. There are two levels to determining falsification; the first is the QA (Quality Assurance) interview where respondents are asked if they have been contacted. If they have not been contacted, a replacement interview is taken. The QA interviewer also records whether he or she feels the interview was fabricated. It is the combination of the respondent and the QA interviewer’s answers that causes a case to replace the original interview. Such cases can be considered as “suspected of falsification.” The second level of determining falsification is that all such replacement cases are investigated by the supervisors. It is the supervisor who determines falsification. When the supervisor determines that an interviewer falsified work, he or she puts all of that interviewer’s work into QA. The way the report is written, it sounds like Figure 4 is meant to show whether this happened or not. If that is the intent, then Figure 4 should show cases determined to be falsified—not cases suspected of falsification.

See comment 10.

The report goes on to state that “the percentage of the replaced interviews that contained errors due to honest interviewer mistakes, to poor respondent recall, or to reasons other than falsification was not measured by the Bureau’s quality assurance program.” This is not true. For all the replacement cases, we determine whether cases are falsified

Appendix I
Comments from the Secretary of Commerce

or whether they fall into any of the categories mentioned above. Therefore, it is measured; however, it appears as the converse of what we report.

See comment 11.
Now on p. 16.

Page 13, paragraph 4— The report states that “when interviews were not complete, the missing data were to be provided through statistical methods, and this would be a source of error in the resulting A.C.E. calculations.” Undoubtedly, imputation leads to some error. However, we have numerous studies over the years showing our imputation schemes to have acceptable error levels. Local variability in quality indicators is unavoidable in any household survey or census, since localities and their populations vary widely. On a regional level, the percent of its geography located in urban areas will have a heavy influence on the quality indicators. For example, the New York region is almost entirely urban, and most of its cases on any survey will be located in urban areas. On the other hand, the Kansas City region has a much smaller percent of its work located in urban areas. For most surveys, comparing those regions would be of little value. In order to assess the extent of “local quality variability,” the comparisons need to be made between demographically and geographically similar areas. For instance, an LCO in Brooklyn, New York, could be compared to a similar LCO in Philadelphia, Pennsylvania, but not to an LCO in rural Nebraska.

See comment 12.
Now on p. 16.

Page 14, paragraph 3— The report states that “local variations in data quality may affect the accuracy of A.C.E. results for some segments of the population.” This is unsubstantiated. We would like to see conclusions that are data-driven. This should be rephrased as a question. The report goes on to state that “42 had to complete over 10 percent of their total caseload with proxy respondents.” This should be 52 LCOs. (Note: in calculating proxy rates, we used the number of proxy interviews divided by the sum of complete and partial interviews. To determine a proxy interview, we use the variables *prxflg* [proxy flag] and *respnum* [respondent number]. If *prxflg* =1 or if *respnum*=99, the case is considered to be a proxy interview.)

See comment 13.

See comment 3.
Now on p. 16.

Page 15, paragraph 1— A sentence should be added to indicate that we calculated a partial interview rate by using the number of partial interviews, divided by the sum of the partials, completes, and noninterviews.

See comment 4.
Now on p. 20.

Page 18, paragraph 3— There was never any intent or requirement to implement operational controls for independence between A.C.E. and census coverage improvement follow-up. This was a conscious decision.

Appendix I
Comments from the Secretary of Commerce

See comment 3.

The clause that reads “1 week after 90-percent completion of nonresponse follow-up in that local census office area” should be revised to read “1 week after 90-percent completion of nonresponse follow-up in all A.C.E. clusters in that local census office area.”

See comment 3.

The final three sentences should be revised to read as follows: “The Assistant Field Division Chief for Evaluation and Research said that exceptions to the start rules had to be approved in headquarters. In fact the only software that would allow an early start to personal visit interviewing was located in headquarters. The regions did not have the ability or the authority for exceptions to be implemented. Any changes required at least Assistant Division Chief level approval.”

See comment 4.
Now on p. 20.

Page 18, paragraph 4— See comments about page 6, paragraph 3. Again, this was a conscious decision. Ad hoc management would have violated our overall management principles.

See comment 12.
Now on p. 23.

Page 19, paragraph 3— The report reads: “If the local census office areas with the worst values of each of these measures have populations that are typically hard to count in the census, these segments of the population may be under-represented in the A.C.E. data, possibly leading to inaccurate reflections of these population segments in A.C.E.-based adjustments.” This is unsubstantiated. We would like to see conclusions that are data-driven. This should be rephrased as a question.

See comment 14.
Now on p. 23.

Page 19, paragraph 4— The report reads: “A relatively small part of the census follow-up work load was not subject to control over its possible overlap with A.C.E. person interviewing; thus, the magnitude of this influence may have been small nationally. However, it may have been significant in some local areas.” This, too, is unsubstantiated and should be rephrased as a question.

The Census Bureau’s Response to the Recommendations for Executive Action

GAO Recommendation: Determine the relationship, if any, between operational measures of person interviewing, such as proxy rates, and the accuracy of A.C.E. estimates of census undercounts as planned.

See comment 15.

Census Bureau Response: The Census Bureau is undertaking very extensive evaluation of the A.C.E. to inform the Executive Steering Committee for A.C.E. Policy for its recommendation regarding uses of A.C.E. results. These evaluations include many cross-tabulations of coverage rates by operational measures of the person interviewing. These evaluations comprise a very extensive analysis of all aspects of the A.C.E. No further evaluations are planned or needed.

Appendix I
Comments from the Secretary of Commerce

See comment 15.

GAO Recommendation: Determine and document the extent, if any, of the actual overlap between census follow-up operations and A.C.E. person interviewing in 2000.

Census Bureau Response: See response to the previous recommendation.

See comment 16.

GAO Recommendation: Determine whether sufficient overlap may have occurred to violate the independence assumption.

Census Bureau Response: The Census Bureau's evaluation on the contamination of census data collected in A.C.E. blocks found, in general, no evidence of contamination of the census due to A.C.E. interviewing. Similar conclusions were reached in several studies on contamination conducted throughout the decade. We see no need for any further work on this topic.

See comment 17.

GAO Recommendation: Determine whether increasing the flow of status data on specific decennial follow-up operations to the managers of independent surveys can help ensure the independence of such surveys, particularly when such operations are scheduled to overlap in the field.

Census Bureau Response: In conjunction with our 2010 census planning, we are beginning an effort to reassess our approach to coverage measurement. As this planning continues, we will be sure to appraise these recommendations with respect to the new approaches to coverage measurement that we are considering.

See comment 17.

GAO Recommendation: Determine what additional steps or controls to preserve the independence of census follow-up and person interviewing, if any, could be implemented for other census follow-up operations that collect enumeration data and are scheduled contemporaneously with person interviewing.

Census Bureau Response: See response to the previous recommendation.

The following are GAO's comments on the Department of Commerce's letter dated October 5, 2001.

GAO Comments

The Bureau generally provided minor technical corrections and additional information. The Bureau also clarified some of our key points and recommendations, which we have reflected in this report and comment on further below.

1. The Bureau noted that figures used throughout the draft report appeared to be inconsistent. We met with Bureau officials and determined that the apparent discrepancies were due to several factors, including the following: (1) our data included Puerto Rico while Bureau data covered only the 50 states, (2) the Bureau initially miscounted the total number of local census offices in its comparison, (3) Bureau results include data from 1,004 quality assurance interviews that replaced the data for the initial interviews, and (4) an error exists in how the Bureau's cost and progress data, upon which we relied, report the number of proxy interviews. In some cases, the Bureau provided us with additional data, and this final report reflects minor changes based on that new information. None of these data changes were significant enough to affect either our conclusions or recommendations. See also comment 13.
2. The Bureau suggested that additional detail be included in figure 1 to indicate that both housing unit matching and person matching operations comprised separate clerical and field follow-up components. We recognize the complexity of those matching operations and will be issuing a separate report on the person matching operation soon. However, to maintain clarity in the figure, we chose not to include such additional detail on the A.C.E. operations that were not the subject of this report.
3. Throughout its response the Bureau suggested various revisions, technical corrections, and clarifications. We revised the report accordingly.
4. The Bureau provided additional information on its decision to permit overlap between census coverage improvement follow-up and the A.C.E. person interviewing operations. The Bureau pointed out that (1) the decision was a conscious one, made in advance of person interviewing, (2) delaying person interviewing until after coverage

improvement follow-up was completed in a local census office area would introduce considerable risk, (3) managing person interviewing on levels of geography below the local census office area would have been impossible, (4) A.C.E. designers believed that the number of coverage improvement cases in an area would not have an effect on A.C.E. data collection, (5) there was never any intent to have operational controls in place between these two operations, and (6) ad hoc procedures without prior headquarters approval were prohibited for A.C.E. We recognize this context, and revised the main body of the report to better reflect it. Nevertheless, while the Bureau's response helps explain why it did not change its procedures or add any new controls or communications when it noticed the increase in the follow-up workload over what was projected, as we note in the report, the increase in workload increased the risk that the independence assumption was violated. There may still be opportunity to implement steps in the future to help ensure the independence of such surveys.

5. The Bureau's St. Louis, Missouri, office should be deleted from the list of Census Bureau regional offices and replaced with Kansas City, Kansas. We revised the text accordingly.
6. The Bureau noted that factors other than the number of people entering their phone numbers on census forms could have accounted for the higher share of person interviewing completed by telephone than was expected. The Bureau suggested that the higher than expected mail return rate of census forms was also a likely factor in the higher telephone interview rate, since this also could have increased the pool of census forms possibly having telephone numbers recorded on them. Our explanation in the draft report was based on interviews with senior Bureau staff in the field division. However, we agree that the mail return also helps explain the higher telephone interviewing rate, and revised the draft accordingly.
7. The Bureau noted that proxy interviews are known indicators of insufficient data quality and not fabrication, as our draft report had suggested. The Bureau noted that indicators of falsification included missing telephone numbers and work days with more than 13 cases. Our draft report was based on language in Bureau training documents for field managers; however, we revised the text to reflect the Bureau comment.

8. The Bureau objected to our comparison of suspected falsification rates in the 2000 Census to those obtained at the three 1998 Dress Rehearsal sites, since the three sites were not representative of the nation. While we agree that the sites are not representative of the nation, and we revised the report to clarify this, we believe that the Dress Rehearsal comparison can provide a reasonable benchmark for regions since, as the Bureau notes, none of the Dress Rehearsal sites were “exceptionally hard-to-enumerate,” and the Bureau believes that hard-to-enumerate areas tend to have higher rates of falsification. We revised the draft to note that half of the regions had falsification rates that fell into the low range of the Dress Rehearsal sites, even though each region contained hard-to-enumerate areas.
9. The Bureau commented that figure 4, which illustrates the regional rates of quality assurance coverage compared to regional rates of suspected falsifications, was “very misleading.” The Bureau said that the figure appeared to attempt to demonstrate whether supervisors were properly following up on cases suspected of falsification. That is not our intent, and our draft did not contain such an implication. We noted in our draft report that the percentage of the interview caseload selected for quality assurance review was expected to vary depending on a number of local circumstances. We reported data on falsification rates only as an example, and because they had been cited as a primary local circumstance during earlier interviews with Bureau staff.
10. The Bureau commented that its quality assurance program *did* in fact measure whether cases contained errors due to honest interviewer mistakes, poor respondent recall, or reasons other than falsification. The Bureau noted that for all replacement cases, it determined whether cases were falsified or fell into the other categories. We revised our report accordingly. However, the quality assurance failure rate that the Bureau calculated and reported includes only those interviews replaced for falsification. We recognize that the quality assurance program was designed primarily to detect falsification, but this definition of “failure” excludes the other sources of rejected interviews and thus understates the rate at which interviews failed to meet Bureau quality standards. After receiving the Bureau’s response, we discussed this with the Associate Director for Decennial at the Bureau, who concurred that the Bureau should consider the broader definition in the future. We have added an additional recommendation for executive action accordingly.

11. The Bureau commented that while imputation for missing data undoubtedly leads to some error, the Bureau had numerous studies over the years showing that its imputation procedures had acceptable error levels. While this explains why the presence of missing data in person interviewing should not by itself be alarming, it does not justify ignoring levels of missing data, the operational quality measures that contribute to missing data, or the methods chosen by the Bureau to deal with missing data. For example, the Bureau recently reported that a variety of alternative statistical models for dealing with missing data gave a wide range of results implying widely varying effects on A.C.E. estimates.⁹ The same report suggested that further research was needed to study these effects.

The Bureau commented that local variability in data quality indicators is unavoidable given variations in local populations and localities. We agree and in our draft report noted that data quality did in fact vary during person interviewing.

The Bureau commented that for most surveys, comparing quality indicators of certain regions would be of little value. The regional comparisons that we made in our report are across the 12 census regions. With the exception of our inclusion of Puerto Rico data with the Boston region data—the census region in which data collection in Puerto Rico was managed and under which census operational data was tabulated in data provided to us by the Bureau—the Bureau reported comparisons of data across the same regions, and of many of the same variables, in a technical memorandum it published in March 2001.

The Bureau suggested that comparisons controlling for demography and geography would be more appropriate to assess the extent of local quality variability. Given the Bureau's acknowledgment that quality variability is unavoidable and that our presentation of local census office area data corroborates that subregional variability exists, we believe that additional comparisons like those suggested by the Bureau are unnecessary to make the general point that variation exists.

⁹Bureau of the Census *Report of the Executive Steering Committee for Accuracy and Coverage Evaluation Policy on Adjustment for Non-Redistricting Uses* (October 17, 2001) p.iv.

12. The Bureau believed that our conclusion linking variations in data quality to possible effects on the accuracy of A.C.E. results was unsubstantiated and suggested wording it as a question. In our draft report we raised the link as a possibility and then recommended that the relationship, if any, be determined between operational measures and the accuracy of A.C.E. estimates. We believe that the conclusion is logical given our prior work linking data quality measures such as missing data rates to errors in A.C.E. results, and provided additional support for making this link.
13. The Bureau noted that there were 52 local census office areas that had to complete over 10 percent of their total caseload with proxy respondents, and not the 42 that we had reported in our draft report. Before receiving the Secretary's response, the Bureau had provided us with additional data. Based on that later data, we counted 49 local census office areas that had to complete over 10 percent of their total caseload with proxy respondents, and we revised the draft text and related figure 5 accordingly. See also comment 1.
14. The Bureau said that our conclusion that the influence of census and A.C.E. overlap may have been significant in some local areas was unsubstantiated. As we noted in the draft report, the Bureau officials we interviewed were unaware of any Bureau attempts to determine the extent of any possible interview overlap in 2000. Such data, if available, might demonstrate whether A.C.E. assumptions were operationally supported in the field. We saw no data on the impact of the overlap that occurred, and we revised the text to state more clearly that the influence of the overlap was a potential one. See also comment 15.
15. The Bureau acknowledged the importance of extensive evaluations of the A.C.E., and referred to the evaluation it is undertaking. We look forward to reviewing this evaluation when it is complete.
16. The Bureau referred to its recent evaluation of the possible contamination of census data collected in A.C.E. blocks, as well as several other similar studies throughout the decade, and saw no need for further work. Some of these studies find weak or only limited indications of contamination of census data in prior censuses, and they all concluded that there was no systemic contamination of census data. The Bureau's most recent evaluation, which is consistent with our recommendation, was released after our audit work was completed. We have revised the draft accordingly.

17. The Bureau said that it was reassessing its approach to coverage measurement. The Bureau gave assurance that it would appraise these recommendations with respect to the approaches under consideration. We look forward to reviewing this appraisal when it is complete.

GAO Contacts and Staff Acknowledgments

GAO Contacts

J. Christopher Mihm, (202) 512-6806
Robert Goldenkoff, (202) 512-2757

Acknowledgments

In addition to those named above, Ty Mitchell, Lynn Wasielewski, Angela Pun, Richard Hung, Janet Keller, Lara Carreon, and staff from our Denver, Los Angeles, Norfolk, and Seattle offices 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