United States General Accounting Office

GAO

Testimony

Before the Subcommittee on the Census
Committee on Government Reform
House of Representatives

For Release on Delivery
Expected at
10:00 a.m. EDT
on Thursday
May 11, 2000

2000 CENSUS

Status of Nonresponse Follow-up and Key Operations

Statement of J. Christopher Mihm
Associate Director
Federal Management and Workforce Issues
General Government Division

GAO/T-GGD/AIMD-00-164

2000 Census: Status of Nonresponse Follow-up and Key Operations

Mr. Chairman, Mrs. Maloney, and Members of the Subcommittee:

I am pleased to be here today to update the Subcommittee on the status of the census. Overall, initial Bureau of the Census data on the conduct of the decennial count is encouraging, with major operations reportedly proceeding on schedule and generally performing as planned.

Particularly noteworthy is the 65 percent initial response rate, which, in matching the response rate to the 1990 Census, surpassed expectations. As we have often noted, although the response rate does not guarantee a successful census, it does reduce cost and scheduling pressures in nonresponse follow-up and subsequent census operations while enhancing data quality. That the Bureau surpassed its expected national response rate goal is a credit to the hard work and dedication of the Bureau's career and temporary employees and of this Subcommittee, which has worked so hard to boost response rates, as well as to the Bureau's government and nongovernmental partners and, of course, the American public.

Currently, the Bureau is engaged in nonresponse follow-up—the largest, most complex, and costly operation of the entire census. My statement today focuses on the progress of the nonresponse follow-up, paying particular attention to the response rate, its impact on the nonresponse follow-up workload, and the Bureau's ability to complete nonresponse follow-up on schedule while maintaining data quality. In addition, I will discuss the Bureau's efforts to redeliver questionnaires initially found to be undeliverable, and the status of the Bureau's data capture operations.

As you know, we have consistently stressed that the census is in many respects a local endeavor because the key ingredients of a successful population count are carried out by locally recruited census employees going from one neighborhood to the next. Likewise, the various enumeration challenges that could reduce the quality of the census often occur locally. Thus, my remarks today are based on interviews with officials from 27 local census offices across the country to obtain their views on the progress of nonresponse follow-up. We selected these offices largely because they had (1) comparatively high nonresponse follow-up workloads, (2) relatively large numbers of hard-to-enumerate groups, and (3) difficulties meeting their temporary employee recruiting goals. We conducted the interviews in early May. To more fully understand nonresponse follow-up operations, we also attended enumerator training at 12 local census offices across the country.

To obtain a national perspective on the status of the census, we analyzed Bureau data, including those data from the Bureau's Census 2000 Management Information System that track the cost and progress of the census at the local census office level. Our analysis included data on the 511 local census offices located in the 50 states.

The Bureau Is Relatively Well-Positioned for Nonresponse Follow-up but Could Face Local Challenges

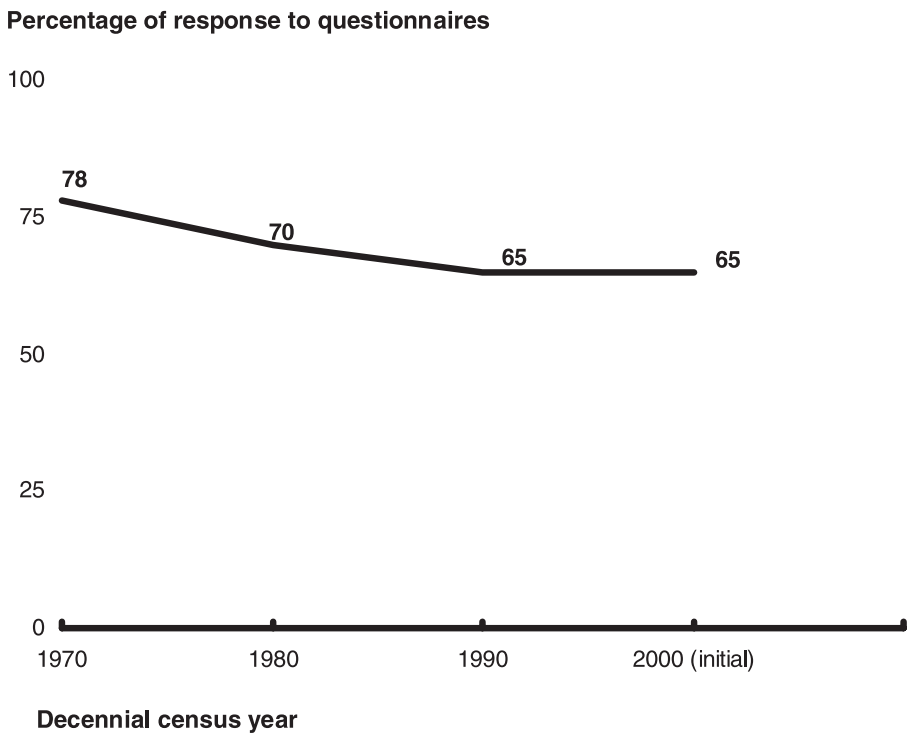
As we have often noted, the Bureau faces a great challenge in completing its nonresponse follow-up workload in the 10-week time frame allotted for it, without compromising data quality. Nationally, the Bureau began nonresponse follow-up in good shape. Because of a higher-than-anticipated mail response rate, the Bureau needs to follow-up with fewer households. At the same time, the Bureau met its staffing needs at most local census offices. Still, some local census offices fell short of their recruiting goals, which could be problematic if they experience significant turnover and need to hire additional employees. Additionally, some local census offices encountered early operational challenges that could affect the productivity and quality of enumerator work.

Bureau Achieved Higher Than Expected National Response Rates

The Bureau achieved an initial response rate of 65 percent as of April 18, 2000, which matched the 1990 rate and exceeded the Bureau's expected national response rate of 61 percent by 4 percentage points.[1] As shown in figure 1, by achieving this 65 percent response rate, the Bureau stopped a three decade long downward trend in census response rates that began when the Bureau first initiated a national mailout/mailback approach in 1970.

[1] For the 2000 Census, the Bureau used what it refers to as an "initial response rate" to provide a measure of the scope of the field follow-up operation with nonresponding households. This initial response rate is defined as the percentage of all questionnaires that are completed and returned by April 18, 2000. The rate includes the number of questionnaires that are mailed back, transmitted via the Internet, or completed over the telephone through the Bureau's Telephone Questionnaire Assistance program. It also includes Be Counted Forms that have census identification numbers.

Figure 1: 2000 Census Ended Downward Trend in Response Rates



Source: U.S. Census Bureau.

With regard to method of response, of the 119 million questionnaires sent by mail or left by enumerators at households, as shown in table 1, most were mailed back. However, a small number of forms were submitted over the Internet and through the Bureau's telephone assistance program.

Table 1: Distribution of Responses by Available Method of Response

Method of response	Number of forms submitted	Percentage of total forms submitted
Mailed back	76,767,689	99.89%
Internet	65,562	.09%
Telephone Questionnaire Assistance	16,814	.02%
Total forms submitted	**76,850,065**	**N/A**

Source: GAO analysis of U.S. Census Bureau data as of April 18, 2000.

The 65 percent response rate is noteworthy given the formidable challenges the Bureau faced in securing public cooperation. Such challenges included attitudinal factors, such as public concern over

privacy and mistrust of government, and demographic factors, such as more complex living arrangements.

The effects of these and related challenges are seen in the continuing problem of bridging the gap between awareness of the census on the one hand and motivation to respond on the other. Various polls have suggested that the public's awareness of the census was high, while, as previously noted, the national response rate was much lower at 65 percent of households. During the 1990 Census, although 93 percent of the public reported being aware of the census, the response rate was 65 percent. Thus, as the Bureau plans for the 2010 Census, it will be important for it to continue to seek approaches that effectively translate the public's awareness of the census into a willingness to respond.

Short- and Long-Form Response Rate Differentials Have Been Increasing

The Bureau also has been unable to close the gap that has existed between questionnaire response rates for the short-form and long-form questionnaires. During the 1990 and 2000 Census cycles, questionnaire response rates were higher for the short-form questionnaire than for the long-form questionnaire, and this gap in response rates has generally widened over time.

For example, as shown in table 2, the differential between the short- and long-form rates ranged from 5.9 percentage points to 8.7 percentage points during the 1988 Dress Rehearsal for the 1990 Census. For the actual 1990 Census, the differential was 6 percentage points. During the Dress Rehearsal for the 2000 Census, the differential ranged from 8.2 percentage points to 14.7 percentage points. While final data are not yet available, the 2000 Census continued with a response rate differential of 12.5 percentage points—over twice that of the 1990 Census.

Table 2: Short- and Long-Form Questionnaire Response Rates, by Census or Dress Rehearsal

Census or Dress Rehearsal	Short form	Long form	Percentage point differential
2000 Census	66.6%	54.1%	12.5%
1998 Dress Rehearsal[a]			
South Carolina	55.4	43.7	11.7
Sacramento	55.4	40.7	14.7
Menominee	40.6	32.4	8.2
1990 Census	66.0	60.0	6.0
1988 Dress Rehearsal			
St. Louis City	50.3	44.4	5.9
East Central Missouri	57.7	52.6	5.1
Eastern Washington	56.5	47.8	8.7

[a]The 1998 Dress Rehearsal was conducted in Sacramento, CA; 11 counties in the Columbia, SC, area; and Menominee County, WI, including the Menominee Indian Reservation.

Source: U.S. Census Bureau.

According to Bureau officials, the Bureau had anticipated a 6.2 percentage response differential between the short and long forms (see table 3).

Table 3: Anticipated and Actual Response Rates to the 2000 Census Short- and Long-Form Questionnaire

Response rate	Short form	Long form	Percentage point differential
Anticipated	62.1%	55.9%	6.2%
Actual	66.6	54.1	12.5
Difference between anticipated and actual	4.5	1.8	6.3

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

However, the actual difference grew to 12.5 percentage points because the response rate to the short form was higher than anticipated, while the response rate to the long form was somewhat lower than anticipated.

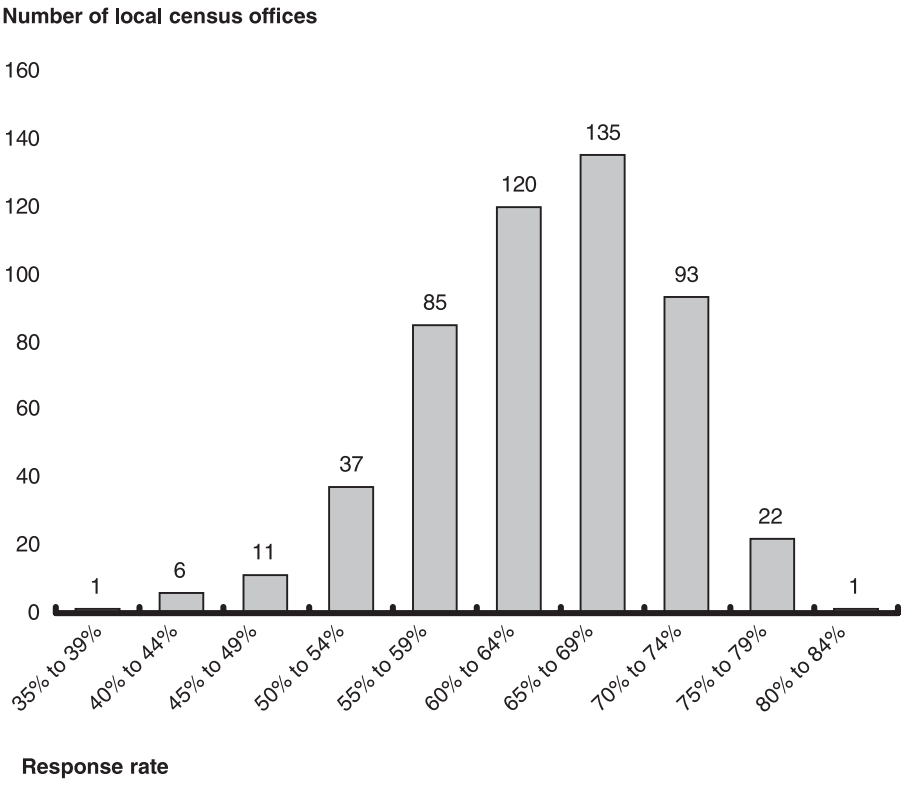
Following the 1990 Census, the Bureau, in expectation of having a more difficult time securing public participation in the 2000 Census, took a number of actions to boost the response rate, including streamlining and simplifying census questionnaires. In our summary assessment of the 1990 census, we noted that developing more user-friendly questionnaires could improve the response rate because it would reduce the time and effort needed to understand and complete a census form.[2] The higher than expected response to the short-form questionnaire suggests that the Bureau's efforts were successful in this regard.

[2] Decennial Census: 1990 Results Show Need for Fundamental Reform (GAO/GGD-92-94, June 9, 1992).

Most Local Census Offices Exceeded Expected Response Rates

Although the 65 percent national response rate provides an overall perspective of the census, local response rates are important because they determine staffing requirements as well as the scope and cost of the Bureau's field follow-up operations. Based on our analysis of Bureau data as of April 18th, we found that response rates at the local census office level ranged from 39 to 80 percent (see figure 2).

Figure 2: Distribution of Initial Response Rates by Local Census Offices



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

According to a senior Bureau official, the Bureau had established expected response rates for four different types of local census offices based on, among other things, the population and housing unit characteristics of the local census offices. Overall, 354 of the 511 local census offices (69 percent) met or exceeded the Bureau's expected response rate. Of the 157 local census offices that did not meet their expected response rate, 125 (80 percent) were the type covering suburban areas, small and medium-sized

cities, towns, and rural areas. This type of local census office makes up 62 percent of all local census offices.

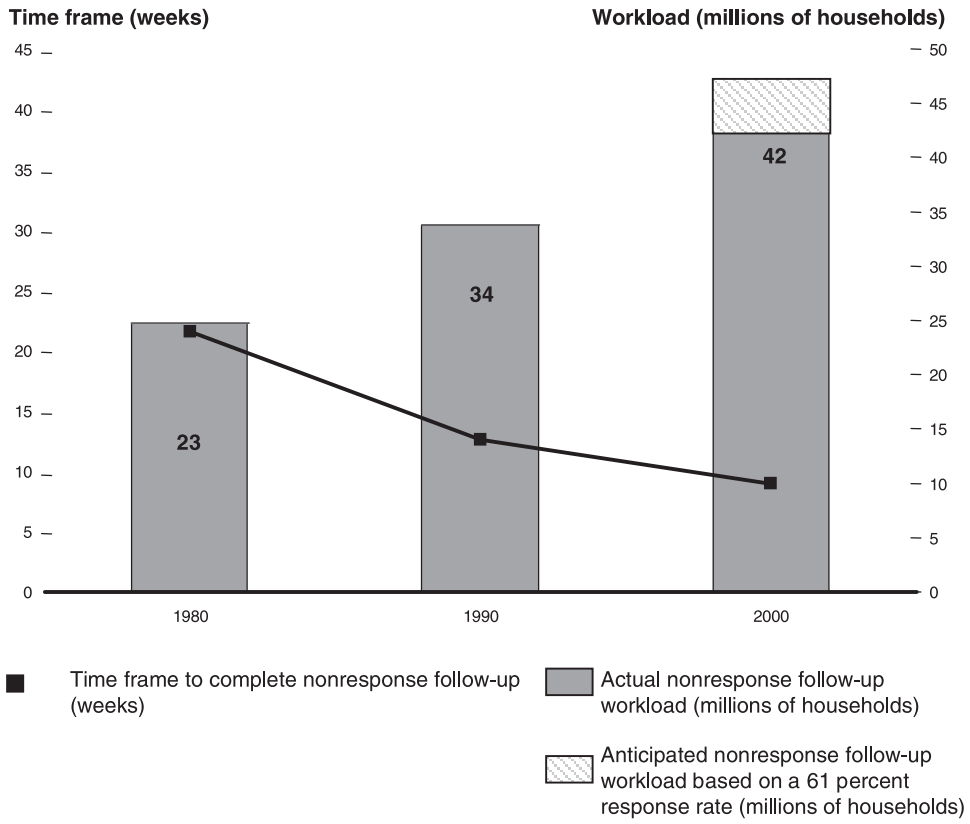
Interestingly, the local census offices covering inner city and urban areas—typically the hardest to enumerate, according to the Bureau—did better than the Bureau expected. Although the Bureau expected they would achieve a 47.5 percent response rate, 92 of the 102 local census offices (90.2 percent) of this type surpassed this rate. This type of local census office represents 20 percent of all local census offices.

The Bureau's Nonresponse Follow-up Workload Is Lower Than Anticipated

The Bureau largely based its schedule, staffing, and funding resources needed for nonresponse follow-up on a 61 percent national response rate. Because the Bureau achieved a 65 percent response rate, the Bureau's actual nonresponse follow-up workload is about 42.4 million housing units-4 million fewer housing units than anticipated.

Still, the Bureau has scheduled only 10 weeks to conduct nonresponse follow-up. So that subsequent operations can proceed on schedule, it will be important for the Bureau to complete the 42.4 million nonresponse follow-up cases within this time frame. However, as shown in figure 3, in the 1980 and 1990 Censuses, the Bureau needed more time to follow up on far fewer housing units.

Figure 3: Nonresponse Follow-up Workload Has Increased Since Previous Censuses, While Time Frames Have Been Compressed



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

The Bureau's ability to maintain this pace will depend on factors ranging from enumerator productivity and turnover rates, to local weather conditions. As of May 7, 2000, data contained in the Bureau's management information system showed that local offices, as a whole, had enumerated 4.3 million (about 10.1 percent) of the nation's 42.4 million nonresponding housing units. It is too early to tell whether this completion rate is indicative of the operation's future progress. However, according to a Bureau official, as of May 8th, the Bureau had completed 17.4 percent of its nonresponse follow-up workload.

The Bureau Met Its National Hiring Goal and Now Must Retain Staffing Levels

To complete nonresponse follow-up on schedule, the Bureau estimated that it would need to fill about 146,000 enumerator positions.[3] Moreover, to address expected turnover, the Bureau planned to "front-load" its workforce by hiring two people for each of these enumerator positions. Data as of May 4th showed that the Bureau hired over 416,000 enumerators. All but 16 of the Bureau's 511 local census offices included in our analysis met or exceeded 90 percent of the front-loaded goal. Assuming that the over 416,000 enumerators are cumulatively working at least the 20 hours per week initially budgeted for the 146,000 positions, which we believe is extremely likely, since the Bureau surpassed its national front-loaded goal as of May 4th.[4]

In addition to hiring a sufficient number of enumerators, officials at most of the local census offices we contacted believe that they have enough bilingual enumerators to follow-up with specific population groups. For example, officials at a local census office in Corpus Christi, TX, said that a high percentage of Hispanics live in the area and that up to half of the office's staff is bilingual. Similarly, an official with a local census office in Chicago said that the office covers a very diverse population, including people of Hispanic, Polish, Chinese, and Lithuanian heritage, and that it has a sufficient number of bilingual staff to conduct follow-up work with each of these population groups.

The Bureau's success in meeting its enumerator hiring goals is due in part to its keeping tabs track of the progress of local staffing efforts and taking quick and sustained action at local census offices that were experiencing recruiting problems. As we noted in our December report, such monitoring of the recruiting process and rapid response to any difficulties would be key to addressing the Bureau's staffing requirements in a tight labor market.[5] For example, during the last 2 weeks of April, the Bureau sent out over 5 million recruiting postcards in targeted areas. Most of these postcard were sent to zip codes in the Boston, Charlotte, and Atlanta census regions, where recruiting was lagging. The Bureau also focused heavily on radio, local television stations, and community newspapers, because these media, along with postcards, generate a quick response

[3] The Bureau adjusted the number of nonresponse follow-up positions in late April to reflect the actual nonresponse follow-up workload.

[4] The Bureau budgeted approximately 200 hours per enumerator position for the nonresponse follow-up operation. Thus, each position is equivalent to about 20 hours per week over the 10 weeks scheduled for this operation.

[5] 2000 Census: Contingency Planning Needed to Address Risks That Pose a Threat to a Successful Census (GAO/GGD-00-6, Dec. 14, 1999).

from applicants. Our recent interviews with local census office managers identified other actions local census offices have taken to boost recruitment, such as increasing the number of recruiting sites, using targeted recruiting ads for specific populations, and developing flexible training schedules for nighttime classes.

The Bureau also responded to recruiting challenges at local census offices by increasing wage rates. Since January 2000, the Bureau has increased wage rates at 31 local census offices. Of these, 11 offices have had their wages increased since we last testified on April 5th. As we noted in our earlier work, higher pay rates helped make the Bureau a more competitive employer when it encountered recruitment difficulties in seasonal resort areas during initial operations for the 2000 Census, as well as in the City of Columbia, SC, during the Dress Rehearsal for 2000.[6]

In addition, the Bureau has continued to work with state governments to obtain exemptions so that individuals receiving Temporary Assistance for Needy Families, Medicaid, and selected other types of public assistance would not have their benefits reduced when earning temporary census income. As you know, we have been supportive of actions that could expand the potential census applicant pool by removing financial disincentives that could discourage people from pursuing census employment. Since we last testified on this issue at the Subcommittee's March 14th hearing, the Bureau obtained exemptions from eight additional state governments. As of April 21, 2000, 44 states and the Virgin Islands had granted an exemption for one or more of these programs.

As previously noted, most local census offices met the Bureau's goal to hire twice as many enumerators as needed to offset expected turnover. To hedge against any additional turnover, the Bureau intends to keep its enumerator positions filled by continuing to hire from its qualified applicant pool.[7] Thus, it will be important for the Bureau to monitor turnover and have a sufficient pool of qualified applicants available to quickly fill any vacancies.

Nationally, the Bureau's pool of qualified applicants stood at over 2.5 million as of April 27h, well in excess of the Bureau's goal of 2.1 million qualified applicants (adjusted from earlier estimates based on the actual

[6] GAO/GGD-00-6; Decennial Census: Preliminary Observations on the Results to Date of the Dress Rehearsal and the Census Bureau's Readiness for 2000 (GAO/T-GGD-98-178, July 30, 1998).

[7] To be counted as qualified, an applicant must pass a basic skills test and a personal background check.

nonresponse follow-up workload).[8] At the local level, 341 of the local census offices 510 (67 percent) had met or exceeded the adjusted recruiting goal as of April 27th. However, 169 local census offices were still short of their recruiting goals, by a total of more than 156,000 qualified applicants. Of these 169 local census offices, 59 fell below their recruiting goal by 20 percentage points or more, and 3 offices had recruited less than half of their adjusted qualified applicant goal.

Thus, nearly 2 weeks into nonresponse follow-up, the Bureau continues to recruit qualified applicants and train them for work on nonresponse follow-up. According to a senior Bureau official, both Bureau headquarters and regional staff will monitor local census offices' production on a daily basis throughout nonresponse follow-up. If some local census offices are unable to meet their production needs, Bureau headquarters and regional staff will work with these offices to take one or more of the following actions: (1) raise pay rates, (2) lower the test score required for selection, or (3) bring in enumerators from neighboring local census offices.

Nonresponse Follow-up Began With Some Early Implementation Challenges

Although the Bureau began nonresponse follow-up in generally good shape nationally, it encountered some early operational challenges. These challenges included a programming error that caused the omission of surname information from nonresponse follow-up address registers, as well as several training and supply glitches.

Bureau Took Actions to Address Surname Problem

As Director Prewitt noted in his April 18th letter to you, Mr. Chairman, the Bureau discovered that, because of a computer programming problem, surnames had been inadvertently omitted from the nonresponse follow-up address registers. According to the Bureau, surname information could help enumerators collect data from intended housing units in situations where questionnaires had been misdelivered in multiunit structures and rural areas with clustered mailboxes.

To remedy the situation, the Bureau decided to produce supplementary address listings that contained surnames, which were to be added to the address registers already produced. Enumerators were then to receive additional training on how to most effectively use the surname address lists. As Director Prewitt stated in his letter, the Bureau expects that this solution will mitigate the problems associated with nonresponse follow-up

[8] The Bureau did not provide recruiting data for the Window Rock, AZ, local census office due to its small nonresponse follow-up workload, and thus we did not include this local census office in our analysis.

materials that do not include surnames. However, he also stated that the listings themselves will be slightly more cumbersome for enumerators to use and thus, may "negatively impact" their efficiency in some cases.

We found that at 8 of the 12 local offices where we attended training, enumerators' training materials did not include a supplementary surname address listing and trainers did not provide training on how to use them. According to a senior Bureau official, prior to the April 24th scheduled training date, the Bureau sent the surname data file and instructions on the use of the surname listings to local census offices and their regional census offices. Regional offices were to inform their respective local census offices about the file sent by the Bureau and share with those offices the instructions sent by the Bureau. According to the Bureau, it sent several electronic messages to the regions alerting them of the availability of the listings and the supplemental instructions. The regions were instructed to make certain that field staff received the materials and understood the procedures. According to the Bureau, as a spot check, several regions were contacted to verify that lists were provided to the enumerators.

Of the 27 local census offices that we contacted following our observations of nonresponse follow-up enumerator training, officials at 24 offices said that the Bureau notified them of the surname problem. They added that the supplemental surname address listings were included in enumerators' nonresponse follow-up address binders at the start of nonresponse follow-up operations and that supervisors provided enumerators information on the use of the supplemental listings. An official at one local census office mentioned that the supplemental lists would add to the already-excessive volume of paper they believed enumerators have to carry while conducting nonresponse follow-up. At another local census office, an official noted that enumerators' manual handling of large volumes of paper may increase their chances of making errors, yet these officials believed that proper supervisory review of enumerators' work should minimize errors.

Enumerator Training Was Generally Well Delivered, but Was Not Always Complete

Based on our observations of nonresponse follow-up training for enumerators at 12 local census offices, we found that trainers at most of these offices were generally prepared and used relevant, "real life" examples of situations that enumerators might encounter and explained how to handle those situations. For example, in a Los Angeles office, trainers discussed how to handle language difficulties, uncooperative residents, and potentially hazardous situations such as vicious dogs.

Nevertheless, at several local census offices, parts of the training were incomplete and key materials were lacking. For example, the trainer at a local census office in Las Cruces, NM, did not show a video on how to conduct an enumeration because, according to the trainer, the office did not have the video that was to be included in the training. At 5 of the 12 local census offices, enumerators did not get a chance to perform a practice enumeration with actual address registers, as was planned for this training, because the address registers were not ready in time. As a result, enumerators missed an opportunity to have on-the-job training and, as a group, discuss their fieldwork experiences prior to conducting nonresponse follow-up on their own.

The lack of on-the-job training could affect enumerators' data collection efforts. For example, on the basis of our observations of enumerator training in a local census office in San Francisco and subsequent discussions with a supervisor and observations of actual enumeration at that office, we found that some enumerators were unsure of how to properly enumerate members of the large transient population prevalent in a neighborhood the census office covered. Had the census workers been able to conduct a practice enumeration exercise in the field prior to going out on their own, it could have helped clarify the enumeration procedures they were to follow.

The Bureau Is Addressing Mailout Questionnaire Delivery Problems

The Census Bureau mailed out about 99 million questionnaires to housing units in mailout/mailback delivery areas of the country. Of these 99 million, the Postal Service was unable to deliver about 11 million questionnaires, which was about 1 million fewer undeliverable questionnaires than the Bureau estimated.

The reasons for these undeliverable-as-addressed (UAA) questionnaires vary. In some cases, housing units that were located within mailout/mailback areas, and that appeared during block canvassing[9] to have mailout/mailback eligible addresses (i.e., street name and building numbers), actually had their mail delivered to post office boxes. The Postal Service generally treated the census questionnaires sent to these addresses as UAA, which is what the Postal Service usually does in such areas for mail addressed to "resident" at a street address. A Bureau official said this is typically what happened in communities, such as Occoquan, VA, that reported that they did not receive census questionnaires. In other cases, housing units were found vacant, addresses had incorrect zip codes,

[9] Block canvassing was a census address listing operation during which census workers canvassed all city-style areas to record addresses.

or whole streets had been renamed since the last time the area had been canvassed by the Bureau.

In anticipation of these problems, the Bureau planned, in consultation with local post offices and Regional Census Offices, to redistribute a portion of the undeliverable questionnaires. Believing that most UAAs would occur in large urban areas, the Bureau planned for UAA redistribution with post offices primarily in those areas, and arranged for those post offices to hold all undeliverable questionnaires until March 18th for Census Bureau employees to pick up and attempt to redeliver. All other post offices were to return their UAA questionnaires directly to the Bureau's National Processing Center in Jeffersonville, IN. As part of nonresponse follow-up, census workers are to conduct census interviews at the addresses for which questionnaires were returned unless an address was already considered questionable and marked as ineligible for the nonresponse follow-up universe.

Preliminary numbers indicate that the Postal Service held about 4.2 million undeliverable questionnaires for over 300 local census offices to attempt to redeliver, and that the Bureau was able to redeliver about 1.6 million of these. And while the Bureau had planned for about 10 million UAAs to be returned to the National Processing Center, as of April 26th, about 9 million had been returned, including those that the Bureau could not successfully redeliver.

Bureau officials have said that,upon hearing reports of clusters of housing units that did not receive questionnaires through either mailout or other delivery methods, they immediately verified whether the reported housing units appeared in the Bureau's master address list. Most of these clusters of missed housing units were in areas without UAA redistribution, and those housing units contained in the Bureau's address list are to be counted during nonresponse follow-up.

To help ensure that any housing unit not already on the Bureau's address list and not returning a questionnaire by other means will still get counted, the Bureau has another procedure in place. During nonresponse follow-up, enumerators are given complete lists of all housing units in their assigned census blocks and are instructed to add and enumerate any housing unit not already appearing on that list. However, according to Bureau officials, since (1) enumerators' primary responsibility is to locate and interview assigned cases, and (2) they are not instructed to recanvass the entire assignment area, enumerators are likely to notice and add such missed housing units only if they are near the housing unit cases already

assigned. Thus, for communities that contact the Bureau about missing questionnaires for housing units that are not on the Bureau's address list, the Bureau is providing information about these specific areas to respective local census offices to help ensure that these areas get added during nonresponse follow-up.

One circumstance that the Bureau cannot remedy with traditional census methods, and that Bureau officials believe is rare, is cases where a housing unit has not been identified by the Bureau's multiple address list-building operations, did not report being missed, and is not identified by an enumerator during nonresponse follow-up. According to a Bureau official, except for remote locations, every area where housing units exist will have been canvassed by Bureau employees at least twice, in addition to any local reviews.

Data Capture Operations and Ongoing DCS 2000 Development Progressing Well

In early April, we testified that each of the Bureau's four data capture centers (DCC) was reporting successful data capture operations and that questionnaires were being processed at a rate that would meet the Bureau's May 26th deadline for completing processing of questionnaires returned by individual respondents (known as mailback processing). Additionally, we noted that delays in ongoing development of new DCS 2000 functions that are needed to capture certain long-form data, if sustained, posed risks to later data capture operations. We attributed these delays to contractor personnel being diverted to address Data Capture System (DCS) 2000 operational problems. At that time, we could not assess other system development risks because plans for completing DCS 2000 development had not been prepared.

We are pleased to report that, as of April 30th, the DCCs were processing questionnaires at a rate that will meet the Bureau's May 26th deadline for completing mailback questionnaire processing. Additionally, the DCS 2000 development contractor has prepared a master plan and adopted an appropriate risk-based approach to modifying DCS 2000's hardware and software configurations, and the contractor is progressing according to its plans. Nevertheless, important development events remain, and the more detailed plans supporting those events have not been finalized.

Data Capture Operations

As of April 30th, each of the four DCCs reports that it has received and checked in more than the expected number of questionnaires. Check-in is the initial step of the data capture process. It entails reading the barcode on each mailed-in questionnaire and sorting the questionnaires for subsequent activities, such as scanning, key from image (KFI), and check out. Similarly, the DCCs report that they are exceeding their respective

goals for the number of forms that have completed the scanning and KFI activities, and although they report that they are slightly below their goal for check out, we do not view this shortfall as significant.

The Bureau has established May 26th as its deadline for completing mailback questionnaire processing, which will entail completing the processing of (1) a backlog of over 13 million questionnaires that have been checked in but not yet checked out and (2) an estimated 2 million yet to be received mailback questionnaires. After this, attention will focus on processing questionnaires completed by enumerators who are collecting data from people who did not return their questionnaires. The processing workloads expected for enumerator forms are substantially lower than the workloads already experienced during the peak of mailback questionnaire processing in late March and early April. Bureau analyses show that sufficient DCC throughput capacity exists to process questionnaires that have been checked in but not yet checked out as well as the mailback questionnaires that the Bureau has yet to receive before its May 26th goal. Based on our review of the Bureau's analysis, we found no reason to question the Bureau's data, calculations, or results.

Moreover, the Bureau reports that DCS 2000's optical character recognition (OCR) accuracy rate was over 99.36 percent at each DCC, exceeding the Bureau's 98-percent accuracy goal. Additionally, the KFI accuracy rate was 97.37 percent or more at each DCC, exceeding the Bureau's 96.5 percent KFI accuracy goal. The KFI keying rate exceeded the Bureau's 5,000 characters per hour goal at each DCC except Jeffersonville, which had a KFI rate of 4,720 characters per hour.

Ongoing DCS 2000 Development

As a result of operational tests at the DCCs, the Bureau realized that the keying rate for KFI was not high enough to meet its master schedule for completing Census 2000 and delivering the apportionment counts by December 31, 2000, as required by law. To resolve this dilemma, the bureau adopted a "two-pass" approach to data capture operations, which required it to modify DCS 2000. During the first pass—from March 6, 2000, until September 2000—the DCCs are capturing only the data necessary to determine the apportionment counts, referred to as 100-percent data.[10]

As we testified in March, the Bureau was creating two configurations or versions of DCS 2000 to enable it to set priorities for data capture operations and thereby meet its deadline for producing apportionment

[10] The 100-percent data are the population and housing information collected for all living quarters in the United States, including the name, sex, and race of each person living in a household.

counts. The first DCS 2000 configuration, designed to support the first pass operations, was completed in early February. This work involved modifying DCS 2000 software to write the images of long-form questionnaires to a mass storage unit and to not present certain data, known as sample data,[11] to keyers. The second DCS 2000 configuration, designed to support the second pass operations, involves modifying the system to retrieve the images of the approximately 22 million long-form questionnaires from the mass storage unit and to present those requiring action to keyers, and then transmitting the resulting data to Bureau headquarters.

Similar to the Year 2000 century date coding changes, the second pass software modifications are not technically difficult to make, but they are pervasive and thus require extensive testing. Specifically, while the second pass changes require changes to fewer than 1000 source lines of code, which is about 1.2 percent of the approximately 85,000 source lines of code in DCS 2000, these change lines are distributed throughout the system. The pervasive nature of the changes thus necessitates extensive analysis and testing to ensure that not only all the changes perform as intended, but also that the changes do have unintended impacts on unmodified code.

The development contractor's approach to making the second pass changes recognizes the pervasiveness of the changes. In particular, the contractor has structured and scheduled development of second pass functionality to provide for extensive testing of the code changes on an incremental basis. Beginning with unit testing of the changes, which has already been accomplished, the evaluation continues with software integration and system integration testing, which are intended to demonstrate that the system meets specified functional requirements. Such an incremental approach to testing is consistent with our published guidance on test management. Additionally, the contractor's schedule contains buffers of time to accommodate changes to the system that test results may necessitate. The schedule also provides for operational testing of the system at the Baltimore DCC, as well as site acceptance testing at each DCC prior to commencing second pass operations, which is intended to show that the system performs as intended in an operational setting.

In addition to extensive testing, the contractor is further minimizing DCS 2000 ongoing development risk by committing a full-time project manager

[11] The sample data include the detailed social, economic, and housing information collected for a sample of living quarters in the United States.

and staff, as opposed to prior plans whereby the development team could be diverted to ongoing DCS 2000 operations. Also, the contractor is taking steps to identify and mitigate program risks. Further, as we noted in our February 2000 report, the contractor is following effective processes for software development.[12] These processes have been independently assessed using the Software Engineering Institute's (SEI) criteria for determining organizations' capability to develop software effectively. SEI's criteria defines five levels of development capability, ranging from level 1 (ad hoc and chaotic) to level 5 (optimized). The contractor, Lockheed Martin Mission Systems, has been independently evaluated as an SEI level 5 development organization.

The Bureau is also taking steps to oversee the contractor's efforts. For example, it is holding weekly technical and schedule status meetings with the contractor, as well as daily project status meetings with the Bureau's data capture program manager. Additionally, the Bureau officials told us that they plan to witness the software integration and system integration tests. Also, the planned operational test will involve the Bureau, the development contractor, and the DCC operations contractor.

Nevertheless, the detailed plans for these various test activities have yet to be developed. The remaining keys to DCS 2000 and second pass future success will thus be the quality of these plans, the plans' effective execution, and the Bureau's close oversight of progress.

In summary, Mr. Chairman and Mrs. Maloney, at this early stage of nonresponse follow-up, the 2000 Census appears to be generally on track. However, as the Bureau recognizes, significant challenges lie ahead. As the Bureau continues its field follow-up efforts, it will be important for it to maintain staffing levels, maximize enumerator productivity, monitor the collection of proxy data, and quickly respond to operational problems.

On behalf of the Subcommittee, we will continue to track the progress that the Bureau and local census offices are making in completing their nonresponse follow-up workload and to monitor the implementation of other census operations.

Mr. Chairman, this concludes my prepared statement. I would be pleased to respond to any questions you or other Members of the Subcommittee may have.

[12] 2000 Census: New Data Capture System Progress and Risks (GAO/AIMD-00-61, Feb. 4, 2000).

Contacts and Acknowledgements

For further information regarding this testimony, please contact J. Christopher Mihm at (202) 512-8676 or Randolph C. Hite on (202) 512-6240.

Individuals making key contributions to this testimony included Jim Bancroft, Thomas Beall, Mark Bird, Richard Burrell, Cristina Chaplain, Betty Clark, Maria Durant, Deborah Eichhorn, Don Ficklin, Robert Goldenkoff, Marshall Hamlett, Richard Hung, Janet Keller, Lily Kim, RoJeanne Liu, Victoria Miller, Vicky L. Miller, Ty Mitchell, Monty Peters, Anne Rhodes-Kline, Miguel Salas, Angela Sanders, Thomas Schulz, Aaron Thorne, Richard Tsuhara, Melvin Thomas, Patrick Ward, and Lynn Wasielewski.