

June 2006

2010 CENSUS

Census Bureau Needs to Take Prompt Actions to Resolve Long-standing and Emerging Address and Mapping Challenges



GAO
Accountability · Integrity · Reliability

Highlights

Highlights of [GAO-06-272](#), a report to congressional committees

Why GAO Did This Study

To conduct a successful census, it is important that the U.S. Census Bureau (Bureau) produce the most complete and accurate address file and maps for 2010. For this review, GAO's specific objectives were to determine the extent to which (1) the Bureau's efforts to modernize the address file and maps are addressing problems experienced during the 2000 Census, (2) the Bureau is managing emerging address file and map issues, (3) the Bureau is able to collect and transmit address and mapping data using mobile computing devices (MCD) equipped with global positioning system (GPS) technology, and (4) the Bureau has a plan to update the address file and maps in areas affected by hurricanes Katrina and Rita. GAO reviewed the Bureau's progress in modernizing both the address file and maps.

What GAO Recommends

GAO recommends that the Secretary of Commerce direct the Bureau to mitigate risks in building its address file and maps. Specific actions include setting firm dates to complete research and evaluations and develop resulting action plans; reevaluating the schedule and staffing workloads for conducting address canvassing; and developing plans to assess resources needed to update the address file and maps along the Gulf Coast. In commenting on a draft of this report, Commerce agreed with each of GAO's three recommendations.

www.gao.gov/cgi-bin/getrpt?GAO-06-272.

To view the full product, including the scope and methodology, click on the link above. For more information, contact Brenda S. Farrell at (202) 512-6806 or farrellb@gao.gov.

2010 CENSUS

Census Bureau Needs to Take Prompt Actions to Resolve Long-standing and Emerging Address and Mapping Challenges

What GAO Found

The Bureau's address and map modernization efforts have progressed in some areas. The Bureau is researching how to correct addresses that were duplicated, missed, deleted, and incorrectly located on maps. However, some deadlines for completing research are not firm, while other deadlines that had been set continue to slip. Thus, whether research will be completed in enough time to allow the Bureau to develop new procedures to improve the 2010 address file is unknown. Also, the Bureau has not fully addressed emerging issues. For one such issue, the Bureau has acknowledged the compressed time frame for completing address canvassing—an operation where census workers walk every street in the country to verify addresses and maps—but has not reevaluated the associated schedule or staffing workloads. Also, the Bureau has allotted only 6 weeks to conduct address canvassing it completed in 18 weeks in 2000 and expanded the operation from urban areas in 2000 to the entire country in 2010.

Mobile Computing Devices for Collecting and Transmitting Field Data



Source: GAO.

Whether the Bureau can collect and transmit address and mapping data using the MCD is unknown. The MCD, tested during 2006 address canvassing, was slow and locked up frequently. Bureau officials said the MCD's performance is an issue, but a new MCD to be developed through a contract awarded in March 2006 will be reliable. However, the MCD will not be tested until the 2008 Dress Rehearsal, and if problems emerge, little time will remain to develop, test, and incorporate refinements. If after the Dress Rehearsal the MCD is found unreliable, the Bureau could face the remote but daunting possibility of reverting to the costly paper-based census of 2000.

Bureau officials do not believe a specific plan is needed to update the addresses and maps for areas affected by the hurricanes. Securing a count is difficult under normal conditions, and existing procedures may be insufficient to update addresses and maps after the hurricanes' destruction—made even more difficult as streets, housing, and population will be in flux.

Contents

Letter		1
	Results In Brief	3
	Background	6
	Uncertainties Surround Completion of Ongoing MAF/TIGER Modernization Research	10
	Emerging Issues Related to Overlapping and Compressed Schedules Pose a Risk to MAF/TIGER Modernization Efforts	18
	Reliability of MCD to Conduct Address Canvassing Activities is Unknown	20
	Bureau Does Not Have a Specific Plan for Updating MAF/TIGER in the Aftermath of Hurricanes Katrina and Rita	22
	Conclusions	24
	Recommendations for Executive Action	25
	Agency Comments and Our Evaluation	26
Appendix I	Scope and Methodology	28
Appendix II	Comments from the Department of Commerce	30
Appendix III	GAO Contact and Staff Acknowledgments	38
	GAO Contact	38
	Acknowledgments	38
Table		
	Table 1: Status of Bureau Efforts to Resolve MAF/TIGER Issues	11
Figures		
	Figure 1: Key Operations Required for a Complete and Accurate MAF/TIGER for the 2000 Census	7
	Figure 2: Group Homes Can Resemble Housing Units	16
	Figure 3: TIGER Map Overlay of an Aerial Photograph	17
	Figure 4: Devastation in the Lower Ninth Ward in New Orleans, Louisiana	22

This is a work of the U.S. government and is not subject to copyright protection in the United States. It may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



United States Government Accountability Office
Washington, DC 20548

June 15, 2006

The Honorable Susan M. Collins
Chairman
The Honorable Joseph I. Lieberman
Ranking Member
Committee on Homeland Security and Governmental Affairs
United States Senate

The Honorable Tom Coburn
Chairman
The Honorable Thomas R. Carper
Ranking Minority Member
Subcommittee on Federal Financial Management, Government
Information, and International Security
Committee on Homeland Security and Governmental Affairs
United States Senate

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

The Honorable Michael R. Turner
Chairman
The Honorable William Lacy Clay
Ranking Minority Member
Subcommittee on Federalism and the Census
Committee on Government Reform
House of Representatives

The decennial census is an important, constitutionally mandated activity undertaken by the federal government that is complex and costly—estimated at \$11.3 billion for the 2010 Census. The data that the census produces are used to reapportion the seats of the U.S. House of Representatives; realign the boundaries of the legislative districts of each state; allocate hundreds of billions of dollars in federal financial assistance; and provide a social, demographic, and economic profile of the nation's people to guide policy decisions at each level of government. The

U.S. Department of Commerce's Census Bureau (Bureau) is responsible for conducting the decennial census, and the success of the census depends in large part on the ability of the Bureau to locate and deliver questionnaires to every person residing in the United States. To successfully accomplish this monumental task, the Bureau must maintain accurate address and map information for every location where a person could reside. During the 2000 Census, Bureau evaluations estimated that of the 116 million housing units in the final census count, about 2.3 million housing units were incorrectly included in the census and about 2.7 million housing units were missed.

One of the Bureau's principal objectives for the 2010 Census is modernizing the Master Address File (MAF)—the Bureau's repository of approximately 130 million addresses to which the Bureau expects to deliver census forms for the 2010 Census. The Bureau also works to ensure the accuracy of the associated mapping system, the Topologically Integrated Geographic Encoding and Referencing (TIGER®) database.¹ The Bureau hopes to improve the completeness and accuracy of MAF/TIGER through a combination of activities, including partnering with state and local governments to verify the address lists and maps and updating maps to reflect correct geographic features. The Bureau will also use satellite-based global positioning system (GPS) technology to correctly locate housing units and door-to-door canvassing to verify the status of all housing units. The combined cost of these efforts is estimated to be about \$536 million (nominal dollars).

An important component of the Bureau's attempts to modernize its address listing and mapping activities will be the planned use of relatively new technology. For the first time, census workers will use a GPS-equipped mobile computing device (MCD) to collect data in the field, including address and map data. The Bureau anticipates that the MCDs will be used in three major census operations, and their successful implementation would allow the Bureau to reduce the amount of paper used, process data in real time, and improve the quality of the information collected.

Because of the critical importance of complete and accurate address lists and maps, under the Comptroller General's statutory authority, we reviewed the Bureau's progress in modernizing both MAF and TIGER. As

¹ TIGER is a registered trademark of the U.S. Census Bureau.

agreed with your offices, we are providing this report to you which contains information that will be useful for your oversight responsibilities of the decennial census. Our specific objectives were to determine the extent to which (1) the Bureau's efforts to modernize the address file and maps are addressing problems experienced during the 2000 Census, (2) the Bureau is managing emerging MAF/TIGER issues, (3) the Bureau is able to collect and transmit address and mapping data using a MCD that is equipped with GPS technology, and (4) the Bureau has a plan to update the address file and maps in areas affected by hurricanes Katrina and Rita.

To meet these objectives, we analyzed relevant evaluations from the 2000 Census and other studies conducted by the Bureau, the Department of Commerce Office of Inspector General, and other organizations. We also reviewed various documents describing the Bureau's MAF/TIGER modernization efforts and interviewed knowledgeable Bureau officials about MAF/TIGER, including the Bureau's plans to update MAF/TIGER in the wake of hurricanes Katrina and Rita. Further, to obtain a firsthand look at how the Bureau's address-building operations and MCDs performed in a real-world environment, we observed address canvassing activities at the 2006 Census Test sites located at the Cheyenne River American Indian Reservation and Tribal Trust Lands in South Dakota and the central portion of Travis County, Texas. Address canvassing is an operation where census workers walk every street in the country, verifying addresses and updating maps. We conducted our work from June 2005 through April 2006 in accordance with generally accepted government auditing standards. Additional information on our scope and methodology appears in appendix I.

Results In Brief

While the Bureau's MAF/TIGER modernization efforts have progressed in a number of areas, it is not clear if research designed to resolve address-related issues from the last census will be completed in sufficient time to improve 2010 address-building activities. During the 2000 Census, the Bureau encountered a number of problems with the MAF, including addresses that were duplicated, missed, deleted, and incorrectly located on the maps. To address those problems, the Bureau has been conducting research and testing some operational changes. For example, the Bureau is researching ways to capture missed addresses for housing units that were hard to find—often associated with apartments in small, multi-unit structures. However, some deadlines for completing research are not firm, while other deadlines that have been set continue to slip. As a result, it is not known whether the research and evaluation efforts currently under way will be completed in sufficient time to allow the Bureau to develop

new methodologies and procedures for improving the MAF by June 2007, the Bureau's announced deadline for baselining all program requirements. In addition, one major research effort using software to identify duplicate addresses (an estimated 1.4 million duplicate addresses were removed during the 2000 Census) did not work any better at identifying true duplicates than what the Bureau already had in place and will not be used in 2010. As a result, duplicate addresses may still be a problem for the 2010 MAF, and if not detected, can result in reduced accuracy and increased cost.

As the Bureau has planned for the 2010 Census, issues surrounding the schedule of address activities have emerged and have not been fully addressed. One such issue revolves around the planning and development of the census amid tight and overlapping schedules for updating addresses and map files. For example, Bureau officials estimate that TIGER maps for 600 to 700 counties of 3,232 counties in the United States will not be updated in time to be part of the local update of census addresses (LUCA)—the Bureau's program to give local, state, and tribal government officials the opportunity to review the address lists and maps and suggest corrections. LUCA participation is important because local knowledge contributes to a more complete and accurate address file, and not having the most current TIGER maps could affect the quality of a local government's review. Also, the Bureau has compressed the time frame for completing address canvassing—an operation where census workers walk every street in the country to verify addresses and update maps. The Bureau has allotted 6 weeks for verifying the nation's inventory of 116 million housing units, although the Bureau took 18 weeks to complete this operation for the 2000 Census. The time to complete address canvassing is a concern because the workload for address canvassing has significantly expanded from including only urban areas in 2000 to including the entire country for 2010. Bureau officials acknowledged the compressed time frame and that, in some areas of the country, bad weather could result in more time being needed to complete address canvassing. Bureau officials did not provide a justification for reducing the amount of time by 12 weeks, but did state that they would need to adjust staffing levels to meet workload demands.

The Bureau's ability to collect and transmit address and mapping data using the MCD is not known. The performance of these devices is crucial to the accurate, timely, and cost-effective completion of address listing, nonresponse follow-up, and coverage measurement activities. During 2006 testing, the MCD used to collect address and map data was slow and locked up frequently. As a result, the Bureau was unable to complete

address canvassing, even with a 10-day extension. Also, some census workers were not always able to get GPS signals for collecting coordinates for housing units. Bureau officials have acknowledged that the MCD's performance is an issue, but believe that a new version of the MCD, to be developed under the Field Data Collection Automation (FDCA) contract awarded on March 30, 2006, will be reliable and functional. However, because the 2008 Dress Rehearsal will be the first time this new MCD will be tested under census-like conditions, it is uncertain how effective that MCD will be, and if problems do emerge, little time will be left for the contractor to develop and test any refinements. Further, if after the dress rehearsal the MCD is found not to be reliable, the Bureau could be faced with the remote, but daunting, possibility of having to revert to a costly paper-based census used in 2000.

Finally, Bureau officials do not believe they need to have a specific plan to update the address and maps files for those areas affected by hurricanes Katrina and Rita. Securing a complete count is difficult under normal circumstances, and the destruction caused by the hurricanes makes it even more difficult because the baseline information the Bureau must work with—streets, housing, and the population itself—will be in flux for some time to come. Bureau officials stated that by 2009, when address fieldwork is set to begin, residents will have decided whether to return to the affected region. Therefore, they believe that by 2009, they will be in a better position to add or delete addresses in the Gulf region affected by the hurricanes. However, Bureau officials could not provide support for the 2009 date, nor have they identified local partners with whom they can monitor this situation. Given the magnitude of the area affected and the degree of destruction, this approach may not be adequate. As a result, the quality of the address and map files could be reduced if the Bureau is not prepared to conduct address operations in those areas affected by hurricanes Katrina and Rita.

In conversations with Bureau officials, it became apparent to us that they are keenly aware of the existing time constraints and challenges detailed in this report. However, the Bureau had not developed risk mitigation plans to address these challenges. Our recommendations, therefore, are intended to make transparent for Bureau and congressional decision makers how those challenges can and should be addressed. At a minimum, the Bureau should have a risk-based mitigation plan in place that includes specific dates for completing research on the address file and an approach for exploring the difficulties the Bureau may face in updating MAF/TIGER along the Gulf Coast. Because time is running short, it is imperative that the Bureau continue to stay focused on identifying and resolving problems

to help ensure that the most accurate and complete address file and maps are produced for the 2010 Census. To facilitate this, we recommend that the Secretary of Commerce direct the Bureau to address methodological, timing, and procedural improvements to building its address file and maps. Specific actions include (1) establishing firm deadlines to complete research, testing, and evaluations of the MAF to prevent missed, deleted, or duplicate addresses and map errors, and develop an action plan that will allow sufficient time to affect the 2010 MAF/TIGER design; (2) reevaluating the 2010 address canvassing schedule in areas affected by bad weather as well as staffing levels to ensure that the status of all housing units are accurately verified throughout the entire country; and (3) developing a plan, prior to the start of LUCA in August 2007, that will assess whether new procedures, additional resources, or local partnerships are needed to update MAF/TIGER along the Gulf Coast for areas affected by hurricanes Katrina and Rita.

On June 2, 2006, the Department of Commerce forwarded written comments from the Bureau on a draft of this report. The Bureau agreed with each of our three recommendations and also noted actions it was taking to address the recommendations. The Bureau's comments also included some technical corrections and suggestions where additional context was needed, and we revised the report to reflect these comments as appropriate. The comments are reprinted in their entirety in appendix II.

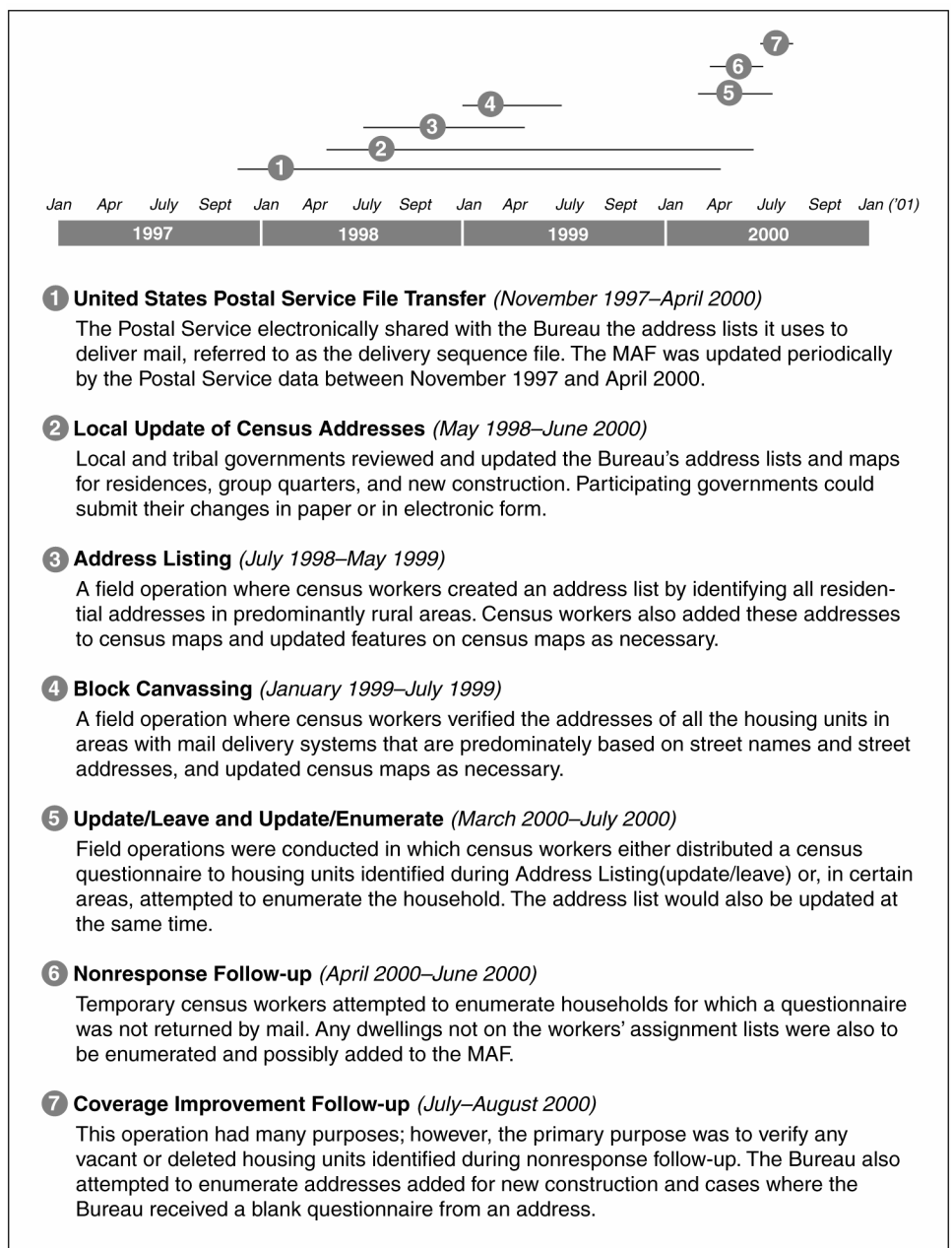
Background

A complete and accurate address list is the cornerstone of a successful census, because it both identifies all households that are to receive a census questionnaire and serves as the control mechanism for following up with households that fail to respond. If the address list is inaccurate, people can be missed, counted more than once, or included in the wrong location. MAF is intended to be a complete and current list of all addresses and locations where people live or could live. The TIGER database is a mapping system that identifies all visible geographic features, such as type and location of streets, housing units, rivers, and railroads. To link these two separate databases, the Bureau assigns every housing unit in the MAF to a specific location in the TIGER, a process called "geocoding."

As shown in figure 1, for the 2000 Census the Bureau's approach to building complete and accurate address lists and maps consisted of a number of labor- and data-intensive operations that sometimes overlapped and were conducted over several years. This effort included partnerships with the U.S. Postal Service and other federal agencies; state, local, and

tribal governments; local planning organizations; the private sector; and nongovernmental entities. The Bureau employed thousands of temporary census workers to walk every street in the country to locate and verify places where people could live. Determining this was no simple task as

Figure 1: Key Operations Required for a Complete and Accurate MAF/TIGER for the 2000 Census



1 United States Postal Service File Transfer (November 1997–April 2000)

The Postal Service electronically shared with the Bureau the address lists it uses to deliver mail, referred to as the delivery sequence file. The MAF was updated periodically by the Postal Service data between November 1997 and April 2000.

2 Local Update of Census Addresses (May 1998–June 2000)

Local and tribal governments reviewed and updated the Bureau's address lists and maps for residences, group quarters, and new construction. Participating governments could submit their changes in paper or in electronic form.

3 Address Listing (July 1998–May 1999)

A field operation where census workers created an address list by identifying all residential addresses in predominantly rural areas. Census workers also added these addresses to census maps and updated features on census maps as necessary.

4 Block Canvassing (January 1999–July 1999)

A field operation where census workers verified the addresses of all the housing units in areas with mail delivery systems that are predominately based on street names and street addresses, and updated census maps as necessary.

5 Update/Leave and Update/Enumerate (March 2000–July 2000)

Field operations were conducted in which census workers either distributed a census questionnaire to housing units identified during Address Listing(update/leave) or, in certain areas, attempted to enumerate the household. The address list would also be updated at the same time.

6 Nonresponse Follow-up (April 2000–June 2000)

Temporary census workers attempted to enumerate households for which a questionnaire was not returned by mail. Any dwellings not on the workers' assignment lists were also to be enumerated and possibly added to the MAF.

7 Coverage Improvement Follow-up (July–August 2000)

This operation had many purposes; however, the primary purpose was to verify any vacant or deleted housing units identified during nonresponse follow-up. The Bureau also attempted to enumerate addresses added for new construction and cases where the Bureau received a blank questionnaire from an address.

Source: GAO compilation of Census Bureau documents.

people can reside in cars, sheds, illegally converted basements and garages, and similar nontraditional and often hidden living arrangements.

For the 2000 Census, the Bureau found that the MAF/TIGER databases were less than complete and accurate. Although the number of errors was small in proportion to the total number of housing units at the national level, the errors could be problematic at lower levels of geography for certain purposes for which census data are used, such as allocating federal assistance to state and local governments.

According to Bureau evaluations conducted after the 2000 Census, the final census count contained approximately 116 million housing units. However, the address file used to conduct the 2000 Census also contained a number of errors.² Bureau evaluations estimate that there were

- 0.7 million duplicate addresses,
- 1.6 million vacant housing units misclassified as occupied,
- 1.4 million housing units not included,
- 1.3 million housing units improperly deleted, and
- 5.6 million housing units incorrectly located on census maps.

In light of these and other problems, the Bureau made enhancing the MAF/TIGER one of three critical components to support the 2010 Census. The other two components are replacing the long form questionnaire with the American Community Survey³ and conducting a short-form-only decennial census that is supported by early research and testing.

For the 2010 Census, the Bureau is making extensive use of contractors to provide a number of mission-critical functions and technologies. One of the technologies to be provided by a contractor is the MCD. Under a contract awarded on March 30, 2006, a new MCD will be developed for the 2008 Dress Rehearsal. To date, the Bureau has tested two models of the MCD—one during the 2004 Census Test and another during the 2006

² The address file used to conduct the decennial census is referred to as the decennial master address file or DMAF. In this report we refer to the address file as the master address file (MAF).

³ The American Community Survey (ACS) will contain the same questions as the long form, but will be mailed monthly to an annual sample of 3 million housing units. With the smaller sample, the ACS is designed to provide the same information at the same level of geographic detail as the long form by means of continuous measurement methodology in which survey responses will be accumulated over time.

Census Test. In January 2005, we reported that the MCD used during the 2004 Census Test to collect nonresponse follow-up data experienced problems transmitting, and the mapping feature was slow. Consistent with our recommendations, the Bureau took steps to improve the dependability of transmissions and correct the speed of the mapping feature.⁴

Due to the critical role of contractors to help carry out the 2010 Census, we conducted a review of major acquisitions for the 2010 Census.⁵ footnote number should not start a line) In that report issued in May 2006, we highlighted the tight time frames the FDCA contractor has for developing and implementing systems to support the upcoming 2008 Dress Rehearsal and recommended that the Bureau ensure that all systems are fully functional and ready to be assessed in time for the Dress Rehearsal. In addition, on March 1, 2006, we testified on the status of the FDCA project.⁶ In that testimony, we discussed the need for the Bureau to

- validate and approve a baseline set of operational requirements for the FDCA contract, because if not, the FDCA project would be at risk of having changes to requirements, potentially affecting its ambitious development and implementation schedule;
- implement an effective risk management process that identifies, prioritizes, and tracks project risks; and
- select detailed performance measures for tracking the contractor's work.

In response to our work, the Bureau stated that they plan to complete these activities as soon as possible.

⁴ GAO, *2010 Census: Basic Design Has Potential, but Remaining Challenges Need Prompt Resolution*, [GAO-05-9](#) (Washington, D.C.: Jan. 12, 2005).

⁵ GAO, *2010 Census: Census Bureau Generally Follows Selected Leading Acquisition Planning Practices, but Continued Management Attention Is Needed to Help Ensure Success*, [GAO-06-277](#) (Washington, D.C.: May 18, 2006).

⁶ GAO, *Census Bureau: Important Activities for Improving Management of Key 2010 Decennial Acquisitions Remain to be Done*, [GAO-06-444T](#) (Washington, D.C.: Mar. 1, 2006).

Uncertainties Surround Completion of Ongoing MAF/TIGER Modernization Research

While the Bureau's MAF/TIGER modernization efforts have progressed in a number of areas, uncertainties and risks remain in dealing with address-related problems that affected the 2000 Census. Currently it is not known whether ongoing research to resolve those problems will be completed in sufficient time to allow the Bureau to develop new methodologies and procedures for improving the MAF by June 2007—the Bureau's announced deadline for baselining all program requirements. One significant cause for this uncertainty is that some deadlines for completing research do not have firm dates, while other deadlines that have been set continue to slip. In addition, one major research effort using software to identify duplicate addresses (an estimated 1.4 million duplicate addresses were removed during the 2000 Census) did not work any better at identifying true duplicates than what the Bureau already had in place and will not be used in 2010. As a result, duplicate addresses may still be a problem for the 2010 MAF, and to the extent they are not detected, can result in reduced accuracy and increased cost.

During the 2000 Census, the Bureau encountered a number of problems with the MAF including (1) missed addresses, where the Bureau failed to include addresses in the MAF; (2) improperly deleted addresses, where the Bureau removed otherwise valid addresses from the MAF; (3) duplicate addresses, with two or more addresses for the same housing unit; and (4) geocoding errors, where addresses were improperly located on a census map.⁷ All of the errors affect the quality of census data. When detected, the errors can increase the cost of the census to the extent they result in rework. Moreover, these errors are associated with a variety of living arrangements and addresses, including small, multi-unit dwellings; dormitories, prisons, and other group living facilities, known collectively as "group quarters," as well as hidden housing units, such as converted basement apartments. As shown in table 1, to address those problems the Bureau has been conducting research and making some operational changes.

⁷ Another type of MAF error identified by the Bureau is misclassifying a housing unit as occupied when it is vacant. However, our focus is on whether an address has been properly captured in the MAF and not the occupancy status of the address. Therefore, we do not discuss occupancy errors in this report.

Table 1: Status of Bureau Efforts to Resolve MAF/TIGER Issues

Types of errors identified in 2000	Primary reasons for those errors	Type of dwelling most likely affected	Actions taken by Bureau	Status of current effort
Missed addresses	Some housing units are difficult to identify.	Small multi-unit structures.	Testing new method to identify clusters of small multi-units.	Testing to be completed by end of 2006.
Improperly deleted addresses	Varied.	Varied.	Tested new method for verifying the status of all housing units marked as deleted in 2006 address canvassing testing.	Evaluation was due January 2006 and that date has been moved to April 2006. The evaluation was not available at the time of this review.
Duplicate addresses	Redundancy and overlap in the address list building process. Address lists were created separately for group quarters and housing units, and some addresses were listed on both lists.	Housing unit with a city-style address. (e.g., 123 Main Street) Group quarters.	Tested address-matching software in 2004. Tested procedures during 2004 and 2006 Census Tests to integrate group quarters and housing unit address lists.	Results indicate matching software is not ready for 2010 Census. 2004 evaluation indicated progress is being made for integrating address lists. Evaluation of 2006 testing was due May 2006 but was not available at the time of this review.
Geocoding errors	Maps not accurate.	Varied.	Collected GPS coordinates for housing units in the 2004 and 2006 tests. Hired contractor to update maps.	2004 test results indicate that workers only used GPS 55 percent of the time. Evaluation was due January 2006 and that date has been moved to April 2006. The evaluation was not available at the time of this review. Contractor updating maps and will be finished in April 2008.

Source: GAO analysis of Census Bureau data.

Research to Identify Hidden Housing Units Is Progressing, but Completion Date Is Uncertain

Although research to find hidden housing units holds promise for a more accurate census, whether the results will be delivered in time to be useful for the 2010 Census is uncertain. While Bureau officials do not have a firm date for completing this research, they do estimate it will be completed by the end of 2006.

According to Bureau evaluations, approximately 1.4 million housing units were missed in the 2000 Census. Missed addresses often result when temporary census workers do not recognize that particular structures, such as tool sheds, are being used as residences. Addresses can also be missed when census workers fail to detect hidden housing units, such as basement apartments, within what appear to be single housing units. This is especially true for urban areas, where row houses have been converted into several different apartments. If an address is not in the MAF, its residents are less likely to be included in the census.

In May 2003, Bureau staff met with the New York City Planning Department to discuss and observe the address problems associated with small multi-unit structures in Queens, New York. After the visit, the Bureau concluded that delivering questionnaires to small multi-unit structures was a problem that needed to be addressed. In response, the Bureau is using the MAF to identify urban areas, including Baltimore, an area west of Chicago, and counties in New Jersey, where small multi-unit dwellings exist, fitting the description of those that were missed.⁸ According to Bureau officials, to accurately identify and count these missed housing units, the Bureau would use update/enumerate procedures—where census workers update the address list and conduct interviews to collect census data—instead of using mailout/mailback procedures, where census forms are mailed to the housing units. Update/enumerate procedures are more labor-intensive and costly than mailout/mailback procedures.

In reviewing the research plan on small multi-unit structures, we found no milestones for completing this research. Bureau officials could not provide a firm completion date, but estimated that the research would be completed by the end of 2006. Without clear milestones for completing this research and action plans based on research results, it is uncertain whether the Bureau will have sufficient time to develop a methodology for identifying all the problematic locations across the country where update/enumerate methodology would be necessary and to inform decision makers on the cost of converting these areas from mailout/mailback procedures to update/enumerate procedures.

⁸ The research project regarding small multiunit structures is also about avoiding duplicate enumerations caused by confusion during mail delivery, as well as follow-up operations.

Research to Prevent Valid Addresses from Being Deleted Is Ongoing, but Completion Date Has Slipped

The Bureau has tested new procedures to validate whether an address initially marked “delete” should be removed from the address file. However, the results from that testing, due January 2006, were delayed until April 2006, and were not available at the time of this review.

For the 2000 Census, the Bureau found that it had mistakenly deleted 1.3 million existing housing units from the address file used to conduct the census. In some instances, this occurred when the Bureau deleted an address that the U.S. Postal Service had coded as a business address, although people were living at that address. According to a Bureau evaluation, when this happens, the Bureau relies on census workers to find and add back those units. Bureau officials stated that identifying residential housing units is difficult for some structures, such as apartments in businesses.

The Bureau would also delete an address if no census form was returned from the unit and if two other census operations determined that the address should be deleted. A Bureau evaluation found that this process identified and removed 8.3 million nonexistent addresses; however, about 653,000⁹ of those addresses were valid and should not have been deleted. The evaluation does not provide an explanation for why these valid addresses were deleted or what could be done in the future to prevent valid addresses from being removed. Concerned that valid addresses were deleted, the Bureau, for the 2006 Census Test of address canvassing, tested a new follow-up quality check procedure designed to verify the status of all addresses that were identified as “delete” during the address canvassing operation. The 2000 Census did have a follow-up operation, but not one specifically for all deleted addresses during the canvassing operation. By building this quality control operation into the address canvassing operation, the Bureau hopes to prevent valid addresses from getting inadvertently deleted. An assessment report of address procedures that were tested in 2005 as part of the 2006 address canvassing operation was to be completed by January 2006. However, the deadline for this assessment slipped until the end of April 2006, and was not available at the time of this review.

⁹ These approximately 653,000 valid addresses that were deleted are a subset of the 1.3 million addresses mistakenly deleted.

Research Efforts on Duplicate Addresses Have Mixed Results

The Bureau has taken actions to prevent duplicate addresses. However, one research effort to identify duplicates using software was found to be ineffective because approximately 10 percent of the time the software would incorrectly identify a valid address as a duplicate address, and as a result, this software will not be used in 2010. According to Commerce officials, it is their philosophy to favor the inclusion of addresses in the census process over the exclusion of addresses. Nevertheless, preventing duplicate addresses in the MAF saves the Bureau from having to make unnecessary and expensive follow-up visits to households already surveyed. Furthermore, preventing duplicate responses also enhances the accuracy of the data.

Bureau studies initially estimated that during the 2000 Census, about 2.4 million duplicate addresses existed in the MAF. The problem was so significant that in the summer of 2000, the Bureau initiated a special follow-up operation¹⁰ to identify and remove duplicate addresses. Research from this special operation confirmed that 1.4 million addresses were duplicates, and the Bureau removed those addresses from the census. However, the operation was not able to determine with certainty whether the remaining 1 million addresses were duplicates. As a result, according to Commerce officials, the 1 million addresses were not removed from the census because those addresses were believed to be a combination of apartment mix-ups and misdelivery of questionnaires, and not duplicates. Had the Bureau identified these 1.4 million housing units before nonresponse follow-up had occurred, it could have saved \$39.7 million (based on our estimate that a 1 percentage point increase in workload could add at least \$34 million in direct salary, benefits, and travel costs to the price tag of nonresponse follow-up).¹¹ Even after the special operation to remove duplicates was completed, the Bureau still estimated that approximately 0.7 million duplicates remained in the MAF in error.

According to Bureau officials, duplicate addresses resulted from the multiple operations used to build the MAF. While the redundancy of having multiple address-building operations helps produce a more

¹⁰ An unduplication operation in the summer of 2000 was implemented to identify and remove duplicate addresses. This operation was not a part of the original 2000 Census plan, but was considered necessary.

¹¹ GAO, *2000 Census: Contingency Planning Needed to Address Risks That Pose a Threat to a Successful Census*, GAO/GGD-00-6 (Washington, D.C.: Dec. 14, 1999).

complete and accurate address list because more opportunities exist for an address to be added to the MAF, any variations in city-style addresses, which are addresses with house numbers and street names, could produce a duplicate entry. For example, the Postal Service, which is the source of many addresses in the MAF, might refer to an address in its database as 123 Waterway Point. A census worker in another address operation might record that address as 123 South Waterway Point. If not detected, two addresses would remain in the MAF for this single residence. To help resolve this problem, in 2004, the Bureau tested whether it could detect duplicate addresses in the MAF by using computerized matching software to link variations in street addresses. In test results, the Bureau found that 90 percent of the potential duplicates identified by the process of “probabilistic matching” were actual duplicates, while 10 percent were valid addresses. Because the number of false duplicates was significantly high, the Bureau decided against incorporating this approach into its plans for 2010 and planned no further testing of the software. As a result of not being able to use this software, duplicate addresses may still be a problem for the 2010 MAF, and duplicate addresses that are not detected can reduce accuracy and increase costs.

At the same time, the Bureau has made some progress toward preventing duplicates. The Bureau is testing new methods to resolve difficulties in distinguishing group quarters (which include dormitories, prisons, group homes, and nursing homes) from housing units, such as single-family homes and apartments. In the 2000 Census, the Bureau used different operations and compiled separate address lists for group quarters and housing units. Group quarters are sometimes difficult for census workers to identify because they often look the same as conventional housing units (see fig. 2). As a result, these homes were sometimes counted twice during the 2000 Census—once as a group quarter and once as a housing unit.

Figure 2: Group Homes Can Resemble Housing Units



Source: GAO, 2004.

One approach to help prevent duplicates that the Bureau tested during the 2004 and 2006 Census Tests is integrating the two address lists and then verifying potential group quarters on that list. Evaluation results from the 2004 testing showed progress was being made for integrating the address lists. The operational assessment report on the 2006 group quarters testing validation/advance visit operation that occurred in 2005, as a part of the address canvassing operation for the 2006 Census Test, was expected by May 30, 2006, and was not available at the time of this review.

Mixed Progress Is Being Made to Properly Identify and Locate Housing Units on TIGER Maps

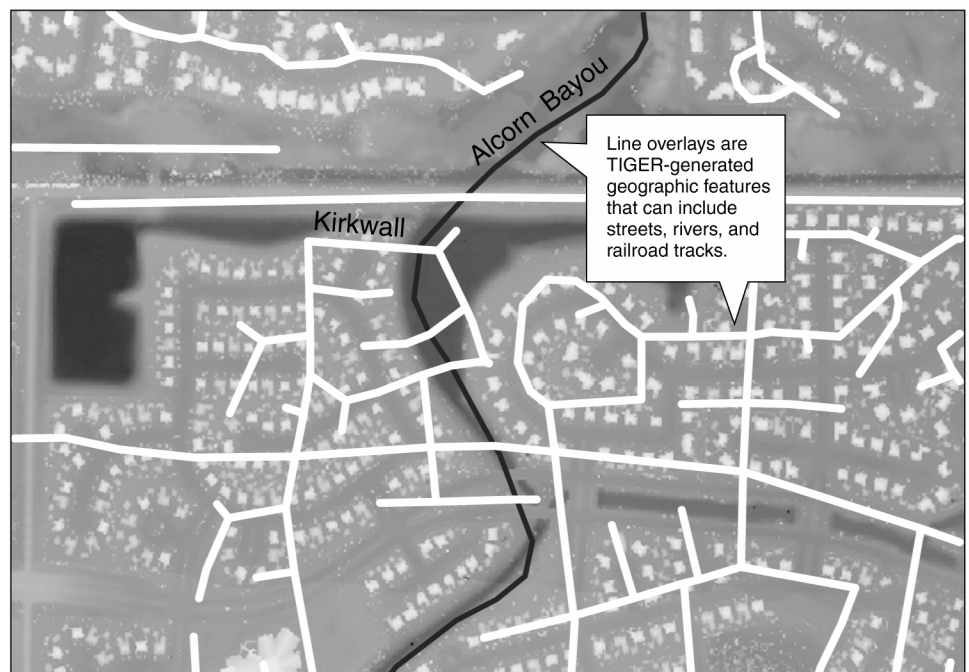
The Bureau is using a contractor to update its TIGER maps and intends to use GPS technology to locate every housing unit across the country precisely. Collectively, these two efforts are designed to avoid the geocoding errors of the 2000 Census, when residences were sometimes counted in the wrong census block. However, progress can be hindered if technical problems associated with the GPS continue.

Bureau evaluations estimated that in 2000, of the nation's approximately 116 million housing units, 5.6 million (about 4.8 percent) housing units in the country were counted in the wrong locations. Resolving geocoding

errors will be important, as census data are used to redraw congressional lines and allocate federal assistance and state funding. For example, in June 2005, we reported that Soledad, California, lost more than \$140,000 in state revenue when a geocoding error caused over 11,000 Soledad residents to be miscounted in two nearby cities.¹²

Geocoding errors are partly attributable to inaccuracies in the TIGER maps that census workers use to verify the locations of residences. As shown in figure 3, roads and other features on TIGER maps did not always reflect their true geographic locations.

Figure 3: TIGER Map Overlay of an Aerial Photograph



Source: U.S. Census Bureau.

To help improve TIGER maps, in June 2002, the Bureau awarded an 8-year, \$200 million contract to correct in TIGER the location of every street, boundary, and other map feature so that they are aligned with their true geographic locations, among other contractual tasks. This work is to be

¹² GAO, *Data Quality: Improvements to Count Correction Efforts Could Produce More Accurate Census Data*, [GAO-05-463](#) (Washington, D.C.: June 20, 2005).

completed on a county-by-county schedule. According to Bureau officials, as of March 2006, nearly 1,700 county maps have been completed, with about another 1,600 to be completed by April 2008.

In conjunction with updating TIGER, the Bureau, as part of its 2010 address canvassing operations, plans to have census workers capture the exact location of every structure on the address list by using GPS receivers. This approach has the potential to resolve the cause of many geocoding errors; however, as we discuss later in this report, when this operation was tested as part of the 2006 Census Test, the GPS receiver did not always operate properly, leaving some housing units without a GPS coordinate to determine their locations. As part of the address canvassing operational assessment report, the Bureau will provide the number and type of map spots collected (GPS, manual, or attached multi-unit). This report, initially due in January 2006, has been delayed and was not available at the time of our review.

Testing GPS coordinates was a part of the 2004 Census Test, and evaluations showed that workers only used the GPS receiver to capture the location of housing units 55 percent of time. The evaluation, however, did not address why census workers did not use the GPS receiver.

Emerging Issues Related to Overlapping and Compressed Schedules Pose a Risk to MAF/TIGER Modernization Efforts

As the Bureau has planned for the 2010 Census, issues surrounding the schedule of address activities have emerged and have not been fully addressed. One key challenge in conducting the 2010 Census is the Bureau's ability to keep the myriad of census activities on track amid tight and overlapping schedules for updating addresses and maps. For example, in planning the various 2010 address list activities, Bureau officials estimate that TIGER maps for 600 to 700 counties (out of 3,232 counties in the United States) will not be updated in time to be part of the local update of census addresses (LUCA)—a program through which the Bureau gives local, state, and tribal government officials the opportunity to review and suggest corrections to the address lists and maps for their jurisdictions.¹³ LUCA is to begin in August 2007, when, according to the current schedule,

¹³ LUCA is an example of how the Bureau partners with external entities, tapping into their knowledge of local populations and housing conditions in order to secure a more complete count. In the Census Address List Improvement Act of 1994, Pub. L. No. 103-430, codified as 13 U.S.C. § 16, Congress required the Bureau to develop a local address review program to give local and tribal governments greater input into the Bureau's address list development process.

the Bureau will still have to update 368 counties in 2008 alone. Because all updates will not have been completed, some counties will not have the most current maps to review, but instead will be given the most recent maps the Bureau has available. According to Bureau officials, some maps have been updated for the American Community Survey, but others have not been updated since the 2000 Census. LUCA participation is important because local knowledge contributes to a more complete and accurate address file. Not having the most current TIGER maps could affect the quality of a local government's review. The Bureau is aware of the overlapping schedules, but officials stated that they need to start LUCA in 2007 in order to complete the operation in time for address canvassing—an operation where census workers walk every street in the country to verify addresses and update maps. Further, Commerce officials stated that the primary focus of the LUCA program is to review and update the address list and not to review and update maps; therefore, not having the improved maps should not affect the ability of LUCA participants to add or make corrections to the census address list. We, however, believe that improved maps would help LUCA participants to provide more accurate address data.

The census schedule will be a challenge for address canvassing in 2010. The Bureau has allotted 6 weeks for census workers to verify the nation's inventory of approximately 116 million housing units. This translates into a completion rate of over 2.75 million housing units every day. The challenge in maintaining this schedule can be seen in the fact that for the 2000 Census, the Bureau took 18 weeks just to canvass "city-style" address areas, which are localities where the U.S. Postal Service uses house-number and street-name addresses for most mail delivery. However, a Bureau official could not explain why the schedule had been shortened by 12 weeks, compared to the 2000 Census.

Although Bureau officials agreed that more time will be needed to conduct the address canvassing operation, especially in the northern sections of the country where bad weather can hinder those operations, they have not reevaluated the schedule. A Bureau official stated that the Bureau would need to assess staffing levels to ensure it will be able to meet workload demands. Meeting the demands of the shortened time frame for completing address canvassing is a concern because the workload for address canvassing has significantly expanded from including only urban areas in 2000 to including the entire country for 2010. Furthermore, in the summer of 2005, when address canvassing was conducted for the 2006 test, the Bureau was unable to finish in 6 weeks because of problems with the new MCD and GPS technology. In its comments to a draft of this

report, Commerce officials said it would work to expand the address canvassing schedule to ensure that it can be done without having a negative impact on other critical decennial activities.

Reliability of MCD to Conduct Address Canvassing Activities is Unknown

The Bureau's ability to collect and transmit address and mapping data using the MCD is not known. The performance of these devices is crucial to the accurate, timely, and cost-effective completion of address listing, nonresponse follow-up, and coverage measurement activities. During 2006 testing, the MCD used to collect address and map data was slow and locked up frequently. As a result, the Bureau was unable to complete address canvassing, even with a 10-day extension. Also, some census workers were not always able to get GPS signals for collecting coordinates for housing units. Bureau officials have acknowledged that the MCD's performance is an issue but believe that a new version of the MCD, to be developed under the Field Data Collection Automation (FDCA) contract awarded on March 30, 2006, will be reliable and functional. However, because the 2008 Dress Rehearsal will be the first time this new MCD will be tested under census-like conditions, it is uncertain how effective that MCD will be, and if problems do emerge, little time will be left for the contractor to develop, test, and incorporate any refinements. Moreover, if after the Dress Rehearsal the MCD is found to be unreliable, the Bureau could be faced with the remote, but daunting possibility of having to revert to the costly paper-based census used in 2000.

Bureau Is Unable to Complete Address Canvassing Operation Because of Technical Difficulties with the MCD

During the address canvassing operation, the technical problems with the MCDs were so significant that the operation did not finish as scheduled.¹⁴ The 6 week operation was expected to run through September 2, 2005, but had to be extended by 10 days (through Sept. 12, 2005). However, the Bureau was still unable to finish the operation, leaving six assignment areas in Travis County, Texas and four assignment areas at the Cheyenne River Reservation, South Dakota not canvassed.

To conduct address canvassing, each MCD was loaded with address information and maps and was also equipped with GPS. Census workers were trained to locate every structure in their assignment area, as well as

¹⁴As noted earlier in this report, during the 2004 Census Test the Bureau also experienced problems with the MCD (different model than the one used in the 2006 Census Test) used to collect nonresponse follow-up data. Specifically, that MCD had difficulties transmitting work and was slow to load maps.

to compare the locations of housing units to address and map data on the MCD and update the data accordingly. They also were instructed to capture each housing unit's GPS coordinates. However, workers we observed and interviewed had problems updating address and map data as well as collecting GPS coordinates, largely because the device's software and GPS receiver were unstable. For example, we observed census workers unable to complete their planned assignments for the day because it took too long to complete address and map updates, as the device was slow to pull up and exit address registers, accept the data entered by the worker, and link a map spot to addresses for multi-unit structures. Furthermore, the devices would often lock up, requiring workers to reboot them.

Census workers also experienced problems with the GPS receiver acquired by the Bureau. Some workers had problems getting a signal, but even when a signal was available, the GPS receiver was slow to locate assignment areas and provide coordinates for map spots. Bureau officials were not certain why the Bureau's equipment was unreliable, but provided several possible explanations: (1) the software, hardware, or both did not function properly, (2) GPS units were not correctly inserted into the device, and (3) too few satellites were available for capturing coordinates. Given the importance of GPS to collecting precise coordinates for housing units, it will be important for the Bureau to understand and correct the source of the problems that affected the reliability of the GPS.

Going into address canvassing, the Bureau was aware that the MCDs had software problems and delayed the address canvassing operation by a month to try to resolve them. The Bureau was unable to resolve the problems, but wanted to test the feasibility of the MCD and decided to go forward with the operation with the goal of learning as much as possible. For the 2008 Dress Rehearsal, the Bureau plans to test a new MCD that is being developed under the FDCA contract. However, less than a year remains for the contractor to develop the MCD that will be used in April 2007 for the canvassing operation of the 2008 Dress Rehearsal. In a May 2006 report,¹⁵ we reported on the tight time frames to develop the MCD and recommended that systems being developed or provided by contractors for the 2010 Census—including the MCD—be fully functional and ready to be assessed as part of the 2008 Dress Rehearsal. In commenting on a draft of this report, Commerce noted that the Bureau

¹⁵ [GAO-06-277](#).

designed the FDCA acquisitions strategy to reduce risks related to cost, schedule and performance, stating that the Bureau required offerors to develop and demonstrate a working prototype for address canvassing. Nevertheless, because the previous two MCD models had performance problems, the introduction of a new MCD adds another level of risk to the success of the 2010 Census.

Bureau Does Not Have a Specific Plan for Updating MAF/TIGER in the Aftermath of Hurricanes Katrina and Rita

The Bureau does not have a plan to update the MAF/TIGER for areas affected by hurricanes Katrina and Rita. On August 29, 2005, Hurricane Katrina devastated the coastal communities of Louisiana, Mississippi, and Alabama. A few weeks later, Hurricane Rita plowed through the border areas of Texas and Louisiana. Damage was widespread. In the wake of Katrina, for example, the Red Cross estimated that nearly 525,000 people were displaced. Their homes were declared uninhabitable, and streets, bridges, and other landmarks were destroyed. Approximately 90,000 square miles were affected overall and, as shown in figure 4, entire communities were obliterated.

Figure 4: Devastation in the Lower Ninth Ward in New Orleans, Louisiana



Source: GAO, 2005.

The task of updating MAF/TIGER for 2010 to reflect these changes will be a formidable one, as much has changed since the 2000 Census. For the 2010 Census, locating housing units and the people who reside in them will be critical to counting the population of places hit by the hurricanes, especially since it is estimated that hundreds of thousands of people

have—either temporarily or permanently—migrated to other areas of the country. To ensure an accurate count, it will be important for the Bureau to have accurate maps and an updated address file for the 2010 Census in those areas affected by hurricanes Katrina and Rita.

Bureau officials do not believe a specific plan is needed to update the address and map files for those areas affected by hurricanes Katrina and Rita. Although Census Day is still several years away, preliminary activities, such as operations for building the MAF, have to occur sooner. Consequently, a key question is whether the Bureau's existing operations are adequate for capturing the dramatic changes to roads and other geographic features caused by the hurricanes, or whether the Bureau needs to develop enhanced or additional procedures before August 2007 when LUCA is scheduled to begin. For example, new housing and street construction in the areas affected by the hurricanes could require more frequent updates of the Bureau's address file and maps. Also, local governments' participation in LUCA might be affected because of the loss of key personnel, information systems, or records needed to verify the Bureau's address lists and maps. Further, the Bureau has not identified local partners with whom it can monitor this situation.

The Bureau's short-term strategy for dealing with the effect of the hurricanes on MAF/TIGER is to see who returns and whether communities decide to rebuild. Bureau officials stated that by 2009, as census workers prepare to go out in the field for address canvassing for the 2010 Census, residents will have decided whether to return to the region. The Bureau believes that by then it will be in a better position to add or delete addresses for areas in the Gulf region affected by the hurricanes. However, Bureau officials could not provide us with information on the basis of their conclusion that by 2009, most affected persons will have made final decisions about whether they are returning to the region. This approach may not be adequate, given the magnitude of the area, population, and infrastructure affected. Therefore, it would be prudent for the Bureau to begin assessing whether new procedures will be necessary, determining whether additional resources may be needed, and identifying whether local partners will be available to assist the Bureau in its effort to update address and map data, as well as in other census-taking activities. In its comments on a draft of this report, Commerce officials stated that there was a team working on how to reflect the impact of the hurricanes in the MAF and that they were aware of the sensitive nature of working with local officials on using data that had not been updated since the catastrophe.

Securing a complete count, a difficult task under normal circumstances, could face additional hurdles along the Gulf Coast, in large part because the baseline the Bureau will be working with—streets, housing, and the population itself—will be in flux for some time to come. According to Bureau officials, different parts of the agency work on hurricane-related issues at different times, but no formal body has been created to deal with the hurricanes' impact on the 2010 Census.

Conclusions

The success of the 2010 Census relies on an accurate and complete MAF, and the Bureau has taken steps to improve the MAF. For example, many of the problems identified in the 2000 Census are being addressed through sequential address list building, the collection of GPS coordinates, and the verifications of deleted addresses. However, several key challenges and sources of uncertainty remain. The management of some of the Bureau's research efforts to resolve problems from the 2000 Census are negatively affected by a lack of specific end dates for that research or because those end dates have slipped. Also, one research effort to prevent duplicate addresses was found to be ineffective and was abandoned altogether. Time to complete research and take the appropriate resulting action is of the essence, as the Bureau has announced that all design features should be baselined by June 2007. If long-standing problems are not resolved, the address file could experience the same problems with missed and incorrectly included housing units as it did in the 2000 Census.

The Bureau must also manage the planning and development of the census amid tight and overlapping schedules. In our view, changing milestones to complete MAF research, the Bureau's tight development schedule for the MCD, and the interdependence of the various address activities could affect the Bureau's ability to develop a fully functional set of address-building operations that can be tested along with other census operations during the 2008 Dress Rehearsal—the Bureau's last opportunity to assess MAF/TIGER under near census-like conditions. If the MCDs do not function as planned in the Dress Rehearsal, little time will remain for the Bureau to develop, test, and incorporate any refinements. This uncertainty places the accuracy and completeness of data collected using the MCD at risk.

Because the MCD has not yet been developed, it will be important for the Bureau to closely monitor the contractor's progress for developing the MCD. In May 2006, we reported on the tight time frames to develop the MCD and recommended that systems being developed or provided by contractors for the 2010 Census—including the MCD—be fully functional

and ready to be assessed as part of the 2008 Dress Rehearsal. However, if after the Dress Rehearsal the MCD is found to be unreliable, the Bureau could be faced with the remote but daunting possibility of having to revert to the costly paper-based census used in 2000.

Finally, the destruction and chaos caused by hurricanes Katrina and Rita underscore the nation's vulnerability to all types of hazards and highlight how important it is for government agencies to consider emergency preparedness and continuity of operations as part of their planning. However, the immediate concern for the 2010 Census is that the Bureau has no plan for how it will successfully update MAF/TIGER in the affected hurricane zone. If problems updating the address file and maps do occur, the census count in those areas affected by hurricanes Katrina and Rita could be inaccurate or incomplete.

In conversations with Bureau officials, it became apparent to us that they are keenly aware of the existing time constraints and challenges detailed above. However, the Bureau had not developed risk mitigation plans to address these challenges. Our recommendations, therefore, are intended to make transparent for Bureau managers and congressional decision makers how those challenges can and should be addressed. At a minimum, the Bureau should have a risk-based mitigation plan in place that includes specific dates for completing research on the address file and an approach for exploring the difficulties that the Bureau may face updating MAF/TIGER along the Gulf Coast. Because time is running short, it is imperative that the Bureau continue to stay focused on identifying and resolving problems to ensure that the most accurate and complete address file and maps are produced for the 2010 Census.

Recommendations for Executive Action

To mitigate potential risks facing the Bureau as it plans for 2010 and to ensure a more complete and accurate address file for the 2010 Census, we recommend that the Secretary of Commerce direct the U.S. Census Bureau to take the following three actions:

- Establish firm deadlines to complete research, testing, and evaluations of the MAF to prevent missed, deleted, and duplicate addresses, as well as map errors, and develop an action plan that will allow sufficient time for the Bureau to revise or establish methodologies and procedures for building the 2010 MAF.
- Reevaluate the 2010 address canvassing schedule in areas affected by bad weather, as well as staffing levels, to ensure that the status of all housing units are accurately verified throughout the entire country.

-
- Develop a plan, prior to the start of LUCA in August 2007, that will assess whether new procedures, additional resources, or local partnerships may be required to update the MAF/TIGER databases for areas along the Gulf Coast affected by hurricanes Katrina and Rita.

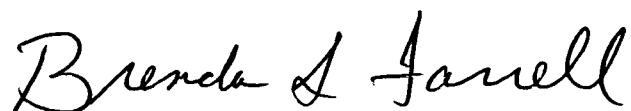
Agency Comments and Our Evaluation

On June 2, 2006, the Department of Commerce forwarded written comments from the Bureau on a draft of this report. The Bureau agreed with each of our three recommendations and also noted actions it was taking to address the recommendations. The Bureau's comments also included some technical corrections and suggestions where additional context was needed, and we revised the report to reflect these comments as appropriate. The comments are reprinted in their entirety in appendix II.

In responding to the first recommendation to develop an action plan that will allow sufficient time to revise or establish methodologies or procedures for building the 2010 MAF, the Bureau stated that it would revise its action plan to reflect final milestones for research to be completed in time for the 2010 Census. Regarding the second recommendation to reevaluate the 2010 address canvassing schedule, as well as its staffing, the Bureau stated that this will be a challenge but that it is committed towards developing a new schedule. Finally, in addressing our third recommendation to develop a plan to assess whether new procedures, additional resources or local partnerships may be required to update the MAF/TIGER databases for areas affected by hurricanes Katrina and Rita, the Bureau stated that it was working on a proposal for additional work in the areas affected by hurricanes Katrina and Rita. The Bureau also noted that conducting additional work will be subject to obtaining funding.

We are sending copies of this report to other interested congressional committees, the Secretary of Commerce, and the Director of the U.S. Census Bureau. Copies will be made available to others upon request. This report will also be available at no charge on GAO's Web site at <http://www.gao.gov>.

Please contact me at (202) 512-6806 or farrellb@gao.gov if you have any questions about this report. Contact points for our Office of Congressional Relations and Public Affairs may be found on the last page of this report. Key contributors to this report are listed in appendix III.

A handwritten signature in black ink that reads "Brenda S. Farrell". The signature is written in a cursive, flowing style.

Brenda S. Farrell, Acting Director
Strategic Issues

Appendix I: Scope and Methodology

To determine the extent to which the Bureau's MAF/TIGER modernization efforts are addressing problems experienced during the 2000 Census, we reviewed pertinent documents, including evaluations of the 2000 Census conducted by GAO, the Bureau, the National Academy of Sciences, and the Department of Commerce's Office of Inspector General. To determine the status of those efforts, we also interviewed cognizant Bureau officials in the Geography Division and Decennial Management Division responsible for implementing the modernization efforts. To assess the extent to which past problems were being addressed, we compared the Bureau's current efforts—including, but not limited to, the 2010 LUCA draft plan, 2004 and 2006 test plans, other research efforts, and TIGER improvement documents—to problems identified in evaluations of the 2000 Census conducted by GAO, the Bureau, the National Academy of Sciences, and the Department of Commerce's Office of Inspector General.

We reviewed the MAF/TIGER contract that was awarded in June 2002 to update the street and geographic features for the TIGER maps, as well as monthly earned-value management system (EVMS) cost and performance reports, to determine whether the deliverable schedule for the contract was on time and on budget. We did not independently verify the accuracy of the data contained in the EVMS cost and performance reports, but we did obtain a certification from the contractor that its EVMS was adequate to provide timely and accurate data from the Defense Logistics Agency.

To determine the extent to which the Bureau is managing emerging MAF/TIGER issues, we focused on planning documents that described proposed 2010 plans. Specific documents we reviewed included the 2010 LUCA draft proposal, 2010 Census decision memorandums, and Bureau papers from National Academy of Sciences and Census Advisory Committee meetings. We also reviewed and compared the timeline for conducting 2000 Census address operations to the proposed plan for conducting 2010 Census address operations. We interviewed officials from the Bureau's Geography Division and the Decennial Management Division on the 2010 plans, 2010 time lines, current status of work, and areas of concern.

To assess the extent to which the Bureau is able to collect and transmit address data using new, GPS-enabled mobile computing devices, we made site visits to census offices on the Cheyenne River Reservation, South Dakota, and in Travis County, Texas, where we observed the address canvassing operation conducted during the summer of 2005 as part of the 2006 Census Test. During these site visits, we also interviewed local and regional census managers and staff, observed address data collection

activities using the MCD, and attended census worker training sessions. We observed and interviewed a total of 38 census workers (16 in South Dakota and 22 in Texas) about the address canvassing operation and the use of the MCD to collect address data. However, the results of these observations are not necessarily representative of the larger universe of census workers. After our visits, we discussed our observations with the Bureau's Technology Management Office, Field Division, Geography Division, and Decennial Management Division.

Finally, to determine the extent to which the Bureau has a plan to update the address file and maps in areas impacted by hurricanes Katrina and Rita, we interviewed Bureau top management officials. Specifically, we discussed whether the Bureau had taken any steps to assess the difficulties it may encounter as it attempts to update the address file and maps and count persons affected by hurricanes Katrina and Rita. We conducted our work from June 2005 through April 2006 in accordance with generally accepted government auditing standards.

Appendix II: Comments from the Department of Commerce



UNITED STATES DEPARTMENT OF COMMERCE
The Under Secretary for Economic Affairs
Washington, D.C. 20230

JUN 12 2006

Ms. Brenda S. Farrell
Acting Director
Strategic Issues
United States Government Accountability Office
Washington, DC 20548

Dear Ms. Farrell:

The U.S. Department of Commerce appreciates the opportunity to comment on the United States Government Accountability Office's draft report entitled *2010 Census: Census Bureau Needs to Take Prompt Actions to Resolve Long-standing and Emerging Address and Mapping Challenges (GAO-06-272)*. I enclose the Department's comments on this report.

Sincerely,

A handwritten signature in black ink that reads "E.R. Anderson".

Elizabeth R. Anderson

Enclosure



Ms. Brenda S. Farrell

2

Commerce Control Number 06-001709

Census Control Number 49503

DMD:TAngueira:PWhite:5/19/06

CQAS:Review:lmh:5/19/06;tdw:5/22/06;kem:5/22/06

CQAS:Final:nth:5/24/06

cc:

ES, US/EA, CQAS (3), A. Moxam, T. Johnson, M. Raines, N. Gordon, R. Swartz, J. Waite,
J. Taylor, T. Mesenbourg, H. Hogan

U.S. Department of Commerce
Comments on the
United States Government Accountability Office
Draft Report Entitled *2010 Census: Census Bureau Needs to Take Prompt Actions to Resolve
Long-standing and Emerging Address and Mapping Challenges (GAO-06-272)*
May 2006

The U.S. Census Bureau generally agrees with the recommendations in this report, but has some concerns and comments about various statements and conclusions.

Regarding the recommendations that begin on page 32:

GAO Recommendation: “Establish firm deadlines to complete research, testing, and evaluations of the MAF to prevent missed, deleted, and duplicate addresses, as well as map errors, and develop an action plan which will allow sufficient time for the Bureau to revise or establish methodologies and procedures for building the 2010 MAF.”

Census Bureau Response: We agree with this recommendation. Throughout this testing cycle, we have maintained a Coverage Improvement Action Plan. In this plan are milestones for conducting research related to improving coverage of housing units in the census. Although this action plan may not be at the level of detail that GAO is suggesting, it has served us well in this decade. We have updated it periodically to reflect results from tests and changes in priorities. Now that the research is winding down, we agree that we should revise this action plan again to reflect final milestones for any research we plan to complete in time for the 2010 Census.

GAO Recommendation: “Reevaluate the 2010 address canvassing schedule in areas affected by bad weather, as well as its staffing, to ensure that the status of all housing units are accurately verified throughout the entire [seems to be a word missing . . . probably ‘country’].”

Census Bureau Response: We agree with this recommendation. We have started working on the scheduling challenges associated with extending the Address Canvassing operation to allow for waves. We are not at the point where we can commit to a new schedule, but we are committed to working toward a new schedule.

GAO Recommendation: “Develop a plan, prior to the start of LUCA in August 2007, that will assess whether new procedures, additional resources, or local partnerships may be required to update the MAF/TIGER databases for areas affected by hurricanes Katrina and Rita.”

Census Bureau Response: We agree with this recommendation. We should have a proposal for any additional work in the areas affected by Hurricanes Katrina and Rita in the near future. Of course, doing additional work will be subject to obtaining additional funding.

Our specific comments and concerns about the report are as follows:

Highlights page - Here, and throughout this report, there are statements about testing the Hand Held Computers (HHCs) in 2005. In order to avoid confusion with our 2005 National Census Test (which did not include any address listing or field interviewing), these statements would be clearer if they referred to this testing as part of the Address Canvassing operation for the 2006 Census Test.

Page 2 - The approximate housing unit count in 2000 was 116 million. Currently, we estimate there will be close to 130 million housing units to be enumerated in the 2010 Census.

Page 4 - The research project regarding small multiunit structures is about avoiding missed units, but is also about duplicate enumerations caused by confusion during mail delivery, during the non-response follow-up (NRFU), or during the coverage follow-up operation. If unit designations are unclear, the Census Bureau will not always be able to determine which questionnaire was delivered to a particular unit, so it will not be certain which units to visit during NRFU. This sort of situation can result in some units being enumerated twice and some not at all.

Page 4 - Concerning the issue of our study of software to identify duplicate addresses, we do not think it is accurate to conclude it did not work. Rather, we found that different types of software did not work any better at identifying true duplicates than what we already have in place. Because of the risks of excluding valid addresses by eliminating duplicates, we will always take a very careful approach to deleting records from the address list based on matching. While it is true that duplicate addresses may still be a problem for 2010, a number of other improvements are tackling that same issue, such as sequential address list operations, Global Positioning System (GPS) coordinates on structures and maps, and integration of the address lists for housing units and group quarters. Additionally, the 1.4 million duplicate addresses deleted from Census 2000 that are referred to did not result entirely from the lack of this type of address matching.

Page 5 - In paragraph 1, the report states that the Census Bureau completed Address Canvassing in 18 weeks in 2000, but on the "Highlights" page up front, it states 8 weeks (which may just be a typo). Overall, it is not clear whether GAO is comparing the 2010 Census Address Canvassing operation to just the Census 2000 Block Canvassing operation which lasted 18 weeks, or to the Census 2000 Address Listing operation which lasted 18 weeks, or to both (which lasted 36 weeks combined).

Page 6 - We will not be identifying housing units as vacant, occupied, or under construction during our 2009 activities. We only will be updating the address list by adding or deleting addresses.

Page 6 - We believe the discussion about updating the maps and lists in the Katrina/Rita areas is not clear. First of all, we need to create contingency plans for any disaster that impacts our list at any time, not just this one. Our research to date on the Katrina/Rita areas has shown that the

address list has been impacted much more severely than the maps. We believe our Address Canvassing operation for the 2010 Census can deal with the impacts on the address file and maps, though of course we need to be prepared for more extensive updating in these areas. In addition, we are aware of the sensitive nature of working with local officials using data that have not been updated since the catastrophe. We are in the process of addressing this problem on a number of levels.

Page 9 - The United States Postal Service (USPS) discussion in this report should refer specifically to their Delivery Sequence File (DSF) because there are many other data products that the USPS provides to us. We received and processed DSFs through April 2000 for Census 2000, including a special extract in February 2000 that included only added units. Later DSF adds were enumerated as part of the Coverage Improvement Followup operation (CIFU) conducted from July 6 through August 22, 2000, so it should be listed as well. In addition, the list should include List/Enumerate areas and remote Alaska. The Local Update of Census Addresses (LUCA) time line includes the dates for the separate additional effort to get newly constructed units added to the list (and enumerated in CIFU), but New Construction was a separate operation that should be mentioned as such in this section if the time line is going to include its dates.

Page 10 - The report sometimes refers to the Master Address File (MAF) and Decennial Master Address File (DMAF) as though they are interchangeable names for the same thing, and we believe this is confusing some of the issues you are discussing. Nothing was deleted from the MAF, and the MAF contains units that were not on the DMAF (which is an extract from the MAF created for conducting the census). When the report discusses estimates of misclassifications, this should refer to the DMAF, not the MAF.

Page 16 - The example given for how the Census Bureau could have mistakenly deleted an address would better be described as an example of how any address could be on the MAF but missing from the census. That is, if the USPS DSF listed a unit as commercial (nonresidential), and no contradictory information was received from any of our other address list development efforts, the address remained on the MAF but was never added to the DMAF. It is not accurate to say that the Census Bureau deleted this type of unit.

Page 17 - The report implies that we did not verify deletes in 2000 and that the 2006 delete verification is new. This is not accurate. For Census 2000, delete verification was conducted during a separate operation which followed the operation that identified the potential delete. For example, addresses classified as delete during Block Canvassing were verified during the 1998 LUCA field verification. Similarly, addresses identified as deletes during NRFU were verified during a separate operation (CIFU), which did not begin until NRFU was completed. With the use of automation, we now are able to identify the deletes quickly and send them back for verification immediately. We believe that doing this concurrently with the operation (in effect, as a Quality Control operation) will improve effectiveness and quality because the verification will happen closer in time to Census Day.

Pages 17 to 20 - The automated detection of duplicates is largely dependent on the amount and quality of the information being matched. Our philosophy has been to “favor” the inclusion of addresses in the census process over the exclusion of addresses. The “unplanned” housing unit unduplication operation that was done in Census 2000 was successful at confirming additional duplicates (1.4 million) in large part because it had the added benefit of decennial person information to consider during the match. For example, an automated process may not be able to absolutely determine if 123 Waterway Point is different from 123 South Waterway Point, but when we include as part of the matching process that a person named John Doe, aged 24, lived at both addresses, we can be more confident in calling it a duplicate address. The report indicates we would save money if this unduplication were done prior to NRFU, but the benefit of the person information is not available before NRFU and, therefore, we could not get similar results if this operation were conducted prior to NRFU.

Page 18 - Partly in response to the lessons we learned about duplication during Census 2000, we are designing our address list development operations to be sequential in 2010. Also, this paragraph fails to mention that we maintained one million cases in the census universe because we believed them to be apartment mix-ups/misdelivery cases (that should not, therefore, be removed from the census), rather than just cases where we were uncertain of their existence.

Pages 23 to 24 - The draft report expresses a concern about some areas not having been completed through the MAF/TIGER Accuracy Improvement Program (MTAIP, which focuses primarily on improving the positional accuracy of the roads in TIGER) prior to the start of the LUCA program: “Because all updates will not have been completed, some counties will not have the most current maps to review but instead will be given the most recent maps the Bureau has available. . . . Not having the most current TIGER maps could affect the quality of a local government’s review.”

The primary focus of the LUCA program is the review and update of the address list, not the review and update of the associated maps. While it is true that the LUCA maps for some communities will not reflect the road positional accuracy improvements and road updates that MTAIP will yield, this in no way interferes with the ability of the affected LUCA participants to add to and correct the census address list. The MTAIP, from its inception, was planned to be completed before the Address Canvassing operation and is on schedule to meet that objective.

Pages 24 to 27 - The statements about the problems we had with the HHC for Address Canvassing do not provide sufficient context. Over a year ago, we stated publicly (on more than one occasion) that we were having difficulty developing the automated instrument and applications for address canvassing. At that time, we also made it clear that we would soon have to decide if enough could be learned by going forward—even with an imperfect instrument—or whether we would have to cancel the test. We later decided to go forward with a goal of learning as much as we could and shared that decision publicly. Therefore, while many of the problems and issues you raise in your report certainly did arise during Address Canvassing, we are disappointed that the report provides none of this background or context for the general reader of the report.

Page 24 (and on Page 5) - The Census Bureau agrees that the six-week period of time initially allotted for conducting Address Canvassing in the 2010 Census was very ambitious. This issue has been under discussion for quite a while, and those discussions continue. As noted in this report, some of the reasons for lengthening the operation are related to staffing and weather concerns, but other reasons relate to the management effort in the regional offices that would be necessary to successfully manage and complete an operation of this size in six weeks.

For example, because Address Canvassing is the first field operation of the decennial census, an expanded period of time for conducting the operation would give the regional offices a critically needed opportunity to train staff, such as area managers and Local Census Office managers, on a flow basis and improve on what was done in previous waves, rather than trying to accomplish this task all at once in a short time frame.

We are committed to working on expanding the length of the Address Canvassing operation and are currently reviewing the schedule to ensure that this can be done without a negative impact on other critical decennial activities. The proposal currently being considered is to conduct Address Canvassing in three six-week waves, similar to Census 2000. The final wave likely would be primarily for areas that are prone to severe weather conditions, which make it difficult for field staff to perform their jobs earlier in the year.

Pages 24 to 27 - Regarding the Field Data Collection Automation (FDCA) contractor's ability to be ready for the Address Canvassing operation, the Census Bureau designed the FDCA acquisition strategy to reduce risks related to cost, schedule, and performance. For example, we required offerors to develop and demonstrate a working prototype for Address Canvassing. The acquisition strategy also provided for a series of technical exchange meetings with each offeror over a period of three months. As a result of these meetings, offerors had the opportunity to learn details associated with the field data collection business processes. The Census Bureau reviewed the completed prototypes in January 2006, prior to contract award.

On March 30, 2006, the Census Bureau awarded the FDCA contract to the Harris Corporation. The Harris Team developed the prototype system with the final solution in mind and plans to refine it, based on prototype lessons learned and based on the requirements decomposition activity that must occur between now and the end of the baseline planning phase in June 2006. The software development for Address Canvassing will occur between June and December 2006. The build cycle includes incremental coding and testing of functionality so that we have visibility into the evolving product, as well as extensive usability and load testing.

Page 29 - As mentioned in the discussion about page 6 above, we will not be identifying vacant, occupied, and under-construction units in advance of the census. Also, we expect that by 2009, the situation on the ground will resemble what will exist in 2010 much more closely than what it would be if we were to make corrections to the list at this point in time. There is, in fact, a team working on the question of how to reflect the impact of the hurricanes in the MAF. Final recommendations should be out soon.

6

Page 30 - Many problems identified in Census 2000 are being tackled on a variety of fronts, from sequential address list-building operations to the collection of GPS coordinates for streets and structures and the recycling of deletes to be validated concurrently with the NRFU operation. This will simplify our efforts to create complete address lists; reduce geocoding error, which contributes to a variety of coverage problems; reduce units deleted in error; and reduce duplication.

Appendix III: GAO Contact and Staff Acknowledgments

GAO Contact

Brenda S. Farrell, (202) 512-6806 or farrellb@gao.gov

Acknowledgments

In addition to the contact named above, Carlos Hazera, Assistant Director; Sheranda Smith Campbell; Betty Clark; Tim DiNapoli; Robert Goldenkoff; Shirley Hwang; Sonya Phillips; Lisa Pearson; Ilona Pesti; and Brendan St. Amant made key contributions to this report.

GAO's Mission

The Government Accountability Office, the audit, evaluation and 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 at no cost is through GAO's Web site (www.gao.gov). Each weekday, GAO posts newly released reports, testimony, and correspondence on its Web site. To have GAO e-mail you a list of newly posted products every afternoon, go to www.gao.gov and select "Subscribe to Updates."

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. Government Accountability Office
441 G Street NW, Room LM
Washington, D.C. 20548

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

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm

E-mail: fraudnet@gao.gov

Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Gloria Jarmon, Managing Director, JarmonG@gao.gov (202) 512-4400
U.S. Government Accountability Office, 441 G Street NW, Room 7125
Washington, D.C. 20548

Public Affairs

Paul Anderson, Managing Director, AndersonP1@gao.gov (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, D.C. 20548