



**PREPARED STATEMENT OF
ROBERT M. GROVES
DIRECTOR
US CENSUS BUREAU**

2010 Census: Master Address File, Issues and Concerns

**Before the Subcommittee on
Information Policy, Census, and National Archives
Committee on Oversight and Government Reform
United States House of Representatives**

21 October 2009

Chairman Clay, Ranking Member McHenry, members of the Subcommittee, I appreciate this opportunity to testify before you regarding the Master Address File for the 2010 Census. When I appeared before the subcommittee on September 22 to discuss my assessment of the 2010 Census, I noted that we would have initial empirical evaluations of the Master Address File at the end of this month. I realize that schedules of the committee did not permit a hearing at that time; I'm happy to tell you everything we know about the quality of the Master Address File as of today. But there is more to learn, and I hope I can meet with the committee when the final analysis is available.

Master Address File (MAF)

The Master Address File, or MAF, is an inventory of all addresses and physical/location descriptions assembled by the Census Bureau, including their geographic locations. The MAF serves as the source of addresses for mailing and

delivering decennial census forms and for physically locating the addresses when necessary (such as during Nonresponse Follow-Up).

In its earliest state, the MAF was an amalgamation of the 1990 Census Address Control File and the United States Postal Service's (USPS) Delivery Sequence File (DSF). Census 2000 was the first major effort undertaken to update and, in a sense, validate the MAF.

Since Census 2000, the Census Bureau has continuously updated the MAF and implemented various enhancements to position it for the 2010 Census. The quality of the MAF is critical to the success of the 2010 Census. The Census Bureau strives to ensure the following tenets of quality:

- Coverage – The MAF should ensure that each housing unit in the United States is included in the Census. Therefore, consistent and complete nationwide coverage is essential to the quality of the MAF.
- Address Completeness - Each address within the MAF should have the most complete information available to ensure delivery of questionnaires and subsequent enumeration of every housing unit and group quarters within the United States. For example, in a multi-unit structure, if we do not have unique unit designations for each apartment, we will likely have problems delivering questionnaires to each address and determining who did not respond so that we can follow up to obtain an interview.
- Spatial Accuracy - The accuracy of the ground location of each address in the MAF ensures that every household is counted in the correct location. The accuracy of this location also ensures the accuracy of counts, statistics, and distribution across statistical and tabulation areas.

Throughout this decade, the Census Bureau has taken various steps to enhance the MAF and our ability to accurately represent the universe of living quarters and the location of those living quarters. For example, we have:

- Adopted an integrated approach to the maintenance of our housing unit inventory and our group quarters' inventory. A key lesson from Census 2000 was that keeping separate address lists for housing units and Group Quarters led to some duplication. While once maintained separately, we have merged our lists into a single source – the MAF.

- Undertaken a multi-year effort to redesign our MAF and Topographically Integrated Geographic Encoding and Reference System, better known as TIGER, databases resulting in a more robust and efficient approach to maintaining and updating the data. For example, prior to the redesign, MAF and TIGER data were stored in county-level files. Now that the data are kept nationally, we can more easily match and update across counties.
- Completed a Nationwide program, the MAF/TIGER Accuracy Improvement Program (MTAIP) where all features (streets, roads, etc.) across the county were spatially aligned. This, along with the collection of GPS coordinates in Address Canvassing will result in a more accurate representation of the location of addresses. This, in turn, will help us successfully locate the addresses during the census and improve accuracy in the tabulation of our data.
- Implemented the MAF Geocoding and Office Resolution (MAFGOR) program, which resulted in updates to the MAF and TIGER that facilitated our ability to determine in which block an address is located.
- Developed a code, referred to as the Address Characteristic Type (ACT), which characterized each geographic block by type of address (city-style, non city-style) and coverage of the USPS's Delivery Sequence File (DSF). This code was used in our 2010 Census planning process to determine the appropriate enumeration methodology. In prior censuses, the determination of enumeration methodology was less precise. We did not use deliverability information from the U.S. Postal Service as part of our criteria to determine where to use the postal service and where to deliver questionnaires ourselves. This resulted in questionnaires for entire neighborhoods being returned to us as undeliverable by the USPS.

In addition to these enhancements, we continuously updated the MAF during the years between the 2000 and 2010 censuses. These efforts provided a current base from which to build the initial universe of addresses for the 2010 Census.

Sources of MAF updates included:

- Updates received twice a year from the USPS including the DSF, as well as other files that facilitate the linkage of one form of an address to another if the addressing formats differ.
- Field work conducted by the Census Bureau field staff updating the address list for specific geographic areas to support the current demographic surveys (such as the Current Population Survey) and to support the Community Address Updating System (CAUS). The CAUS aims to incorporate changes to addresses in non-urban communities in the

U.S. These areas are excluded from automated updates from Postal Service files and require the use of trained field staff to add, update, or delete non-city-style addresses in the MAF.

- Address and location (geocoding) information from the American Community Survey.

Collectively, the update of the MAF from various partnership and field operations and the verification of these addresses through the 2010 Census Address Canvassing are the first steps to ensuring the quality and coverage of addresses in the 2010 Census. But our efforts do not end there.

Local Update of Census Addresses (LUCA) Program

While we implemented these ongoing updates to the MAF, we also implemented various efforts specific to enhancing it for the 2010 Census. A major enhancement over our approach for Census 2000 involved the Local Update of Census Addresses (LUCA) program.

LUCA plays a critical role in ensuring the accuracy and success of the 2010 Census. By working with tribal, state, and local governments, we incorporate updates and new information into our MAF and digital mapping system, TIGER that are based on their timely and intimate local knowledge of these local governments. This is a key program to enhance the transparency and improve the quality of the decennial census.

Mr. Chairman, I want to thank you, along with Representative Maloney, the former ranking member of this subcommittee, Representative Michael Turner, Representative Michael Thompson, and the former ranking member of the Full Committee, Representative Tom Davis for your support of the LUCA program last year. The "Dear Colleague" you coauthored on March 28, 2008, helped educate members of Congress about the LUCA program and stressed the importance of local government participation. Your support helped get the word out about LUCA registration and increase participation in the program.

The Census Bureau kicked off the LUCA program in 2007. We invited 39,379 tribal, state and local governments to participate in the program. Approximately 11,500 governments registered, about 29 percent of the governments eligible to participate, but representing approximately 92% of housing units. In Census 2000, more than 18,000 governments registered for the Census LUCA Program, representing 92 percent of the housing units in

Mailout/Mailback areas and representing 68 percent of the housing units in the rest of the country; the drop in registrations is in some measure attributable to higher-level governments, including states, registering to participate on behalf of lower-level entities within their boundaries. Governments were offered three options for participating: options 1 & 2 involved the sharing and review of confidential Census Address data with feedback; option 3 provided for the sharing of non-confidential Census Address Counts for governments to review without feedback.

All three options allowed governments to participate according to their needs and resources. They could choose to focus on areas where addresses are likely to be missed, such as areas of new construction, areas changed from single family to multi-family units, commercial areas that were converted to residential, and areas undergoing significant change due to economic displacement or natural disasters. In addition, participants had 120 days to review the Census Address List, as opposed to the 90 days afforded in 2000.

We received submissions reflecting changes to the address list and/or count list and/or maps from 79 percent of the governments that registered to participate, or just over 8,100 participants. This compares to 67 percent of LUCA 1998 participants and 48 percent of LUCA 1999 participants who provided updates, and may indicate that the Census Bureau better communicated the LUCA participation requirements for the 2010 Census. Submissions from LUCA participants were matched against the MAF, and then sent out in the Address Canvassing operation for verification. The Geography Division has now processed the results of Address Canvassing, including those LUCA submissions that were verified in the field, into the MAF and TIGER databases, and LUCA feedback materials are being produced.

There are several reasons that the 2010 LUCA program might be more effective than that of the prior decade.

2010 LUCA Program improvements

- Single cycle of review for all addresses (city and non-city style) rather than separate cycles that were especially confusing to governments with both types of addresses in Census 2000 LUCA. These governments received their lists of city style and non-city style addresses separately and at different times, which made it difficult to

determine what addresses were actually missing from the list altogether.

- 120-day review period for participants rather than 90 days allowed for Census 2000 LUCA. Although there were significant delays in the deployment of LUCA materials to participants, the 120-day review period was preserved for all but about 30 participants (one quarter of one percent).
- Multiple program participation options that for the first time allowed governments to submit their own address lists for the Census Bureau to use for matching/update, and a non-Title 13 confidential option for governments unwilling/unable to meet our data confidentiality and security requirements.
- Free availability of easy-to-use desktop software to assist participants with their review and update of the census address list and maps.
- Significantly earlier communication with eligible governments about the LUCA Program to allow them more preparation time; this was accomplished with an informational mailout to eligible governments 6 months prior to the LUCA invitation mailout in August, 2007. Contact was made with the state affiliates of the National Association of Counties and the National League of Cities, constituting 44 State associations and 49 State leagues, throughout the spring and summer of 2007.
- Encouragement for different levels of government to collaborate on participation, which helped overcome participation barriers facing under-resourced governments.
- State governments' eligibility to participate in LUCA for the first time, with 28 registering to do so.
- Computer-based training for the submission process available for the first time.

The preliminary figures from our LUCA program show:

- about 8 million addresses were provisionally added to the MAF for verification in Address Canvassing;
- 30 million LUCA submissions matched to addresses already [?] on the MAF;
- 2 million corrections to addresses on the MAF;

As planned, we requested and received updates from LUCA participants in advance of the 2010 Census Address Canvassing operation.

Local Update of Census Addresses Appeals Process

Beginning in October 2009 and ending December 2009, each Option 1 and Option 2 LUCA participant that either provided updates or certified that the list was correct receives a feedback packet from the Census Bureau. Option 3 participants may not appeal, because they do not receive the detailed address level feedback materials required for the basis of an appeal, nor do they comply with Title 13 confidentiality requirements. The feedback packet contains the most recently updated full address list as well as specific feedback on action taken by the Census Bureau for each LUCA add, correction or block count challenge. These packets are sent on a rolling basis; the first feedback packets were sent last week and the last will be sent by December 11, 2009. Participants that disagree with the Census Bureau's determination on any of the corrections they offered or any address deletions from the Address Canvassing Operation have 30 days from the time they receive the feedback to appeal each such determination to the 2010 Census LUCA Appeals Staff. We will know the universe of governments that filed an appeal by late January. The 2010 Decennial Census LUCA Appeals Staff, a temporary Federal entity independent of the Census Bureau and overseen by the Office of Management and Budget (OMB), will review the evidence provided for each appealed address or challenged block. Once the Appeals Officer concludes his/her review and submits a recommendation, the Director of the Appeals Staff will submit his determination, which is final, to both the Census Bureau and the eligible government. An explanation for the determination will be included. All appeals must be processed and resolved by Census Day, April 1, 2010. Addresses for which the appeal is sustained will be visited by Census field staff and enumerated if found during the Vacant Delete Check Operation.

In Census 2000, in our Mailout/Mailback areas, 645 governments appealed 313,853 addresses; the Appeals Office accepted 303,410 of them; and ultimately 141,580 of them were enumerated in the census. In the rest of the country, 620 governments appealed 18,442 addresses; the Appeals Office accepted all of them; and ultimately 10,053 of them were enumerated in the census.

Address Canvassing

The first nationwide field operation to update the MAF in the 2010 Cycle was Address Canvassing, conducted in the summer of 2009. During Address Canvassing, field staff performed a 100-percent canvassing operation to validate all existing addresses on the MAF (including those from early in the decade and from the LUCA program) and to add any newly discovered addresses. This operation was conducted in all areas of the country except remote areas in Alaska, approximately 28,000 addresses, and Maine, approximately 7,000 addresses, out of the approximately 134 million addresses in the country.

To ensure the quality of the address list resulting from the 2010 Address Canvassing operation, we designed and successfully implemented several measures during the operation:

- Address-level edits built into the software to ensure data quality at the time of collection.
- Development of special job aids to address small multi-unit structures, hard-to-locate units, and areas impacted by natural disasters.
- Validation of addresses added to the address list through non-field programs such as LUCA or DSF updates.
- A quality control component in the operation.
- Verification of deleted addresses during Address Canvassing.

Additionally, before applying any of the updates from the Address Canvassing operation to the MAF, we also instituted several components of quality assurance into the MAF updating processing, including:

- Validation checks and acceptance criteria on incoming data.
- Matching of addresses to existing data in the MAF.

We successfully completed the Address Canvassing operation over the summer, whereby census staff checked a total of 145 million addresses, making additions or deletions where necessary. As previously stated, this included 8 million addresses added by tribal, state and municipal governments in the Local Update of Census Addresses program, approximately 3 million more addresses than 2000. At this time, we are analyzing the characteristics of the MAF.

Preliminary evaluation of the Address Canvassing Operation indicates it was successful in that it was produced on time and within the required quality

assurance parameters. Listers also updated and verified existing addresses, added addresses not currently on the address list and deleted addresses from the list that were not found or existed in another form. Listers updated maps by deleting and adding features, and updating feature names.

We do have some initial empirical assessments. These early figures show that about 21 million addresses fall into the deletes, duplicates, or non-residential category. Of these, 16 million were deletes (53 percent were from Census 2000 lists; 27 percent from LUCA; 7 percent from USPS; and 12 percent from other sources). About 4 million were duplicates, and one million were non-residential, meaning they were determined during address canvassing to be businesses or structures for other uses.

Other early figures show that 98 million addresses were verified as is, 20 million were corrected (e.g., street name change), 5 million were moved to another block and 10 million were added (not on address list before and not from LUCA). The current address list stands at 134 million addresses.

About 2 million were coded as Other Living Quarters. By “other living quarters” we mean structures where larger numbers of unrelated individuals live (such as dormitories and assisted living quarters) as well as trailers or other movable units. As I speak, we’re re-inspecting these structures in the newly instituted Group Quarters Validation operation.

With respect to the results of Address Canvassing on LUCA updates, our initial results show that 66 percent of the provisionally added LUCA addresses were deleted, identified as a duplicate, or found to be nonresidential. About 29 percent of the provisional addresses were verified, corrected, or moved; 5 percent were unresolved in Address Canvassing and will remain in the enumeration universe.

Review of the entire Address Canvassing Operation is now underway. When this process has concluded and the final summaries of how many changes were made to the Master Address File are available, I will be happy to provide them to the Subcommittee.

At this point, we can say that approximately 2 million addresses have been sent to the Group Quarters Validation Operation and 132.4 million addresses are considered to be housing units at this point. Of the known housing units, 119.1

million are in Mailout/Mailback areas, 11.9 million are in Update/Leave areas, 1.4 million are in Update/Enumerate areas, and 27,600 are in Remote Alaska.

New Construction Program

The New Construction program will give governments the opportunity to submit city-style addresses for units constructed after the Address Canvassing operation. The program is offered to local and tribal jurisdictions that contain blocks where the Census Bureau plans to mail questionnaires to the housing units. In other areas, the Census Bureau will have enumerators hand deliver questionnaires to all housing units in each block and update the address list for any new housing units. Tribal and local governments with any area included in mailout-mailback are eligible to participate whether or not they participated in the LUCA Program. Although state governments were eligible to participate in the LUCA, they are not included in the New Construction Program, primarily because they do not have on-the-ground knowledge about construction activity at the local level.

The Census Bureau will send the program materials to the New Construction liaison. The New Construction liaison must submit a list of city-style addresses, assigned to the census blocks within its jurisdiction in the Census Bureau predefined format. The maps or spatial data are for use as a reference for assigning Census tract and block codes (geo-coding) for each submitted address. No street or boundary updates will be accepted. Address lists submitted without geo-coding information will not be accepted.

We invited 28,684 governments to participate in the New Construction Program. Of them, thus far we heard back from 14,528 governments, with 5,882 governments having registered for the program by the October 8, 2009 deadline.

Ongoing 2010 Census MAF Updates

As was mentioned previously, one critical indicator of the quality of the address list is how well it covers the housing unit universe. We need to ensure that every living quarter is represented on the MAF once and only once. We also need to ensure that the address information we have is sufficient to either mail a questionnaire to a specific address and/or for us to locate an address to drop off a questionnaire or revisit the address in our various 2010 Census operations.

Although we have concluded the Address Canvassing operation, updating the MAF does not end with Address Canvassing. Many additional programs and operations will impact/improve the coverage of the MAF. These operations and programs include:

- Additional updates from the US Postal Service's Delivery Sequence File (DSF) - In October 2009 and again in February 2010, the Census Bureau geography staff will match addresses in the DSF to the MAF. During this match, we will harvest city-style addresses that are not currently on our MAF. This match occurs in areas that are considered mail-out/mail-back. The updates from the October 2009 DSF will be included in the initial Census mail-out, and updates from February 2010 will be included later in the Census enumeration process (e.g., during nonresponse followup), but will not receive a mail-out questionnaire.
- The Count Review Program - In early 2010, the Count Review Program will provide the Federal State Cooperative Program for Population Estimates (FSPCE) an opportunity to provide housing unit data for comparison to the extract of the MAF used for enumeration. The goal is to identify clusters of housing units that are missing from the enumeration universe. Additional review provided by the Population Division will ensure a consistent minimum review across the nation.
- Additions, updates, and deletions from other field operations - Within most major 2010 Census field operations after Address Canvassing (Group Quarters Validation, Update Leave, Update Enumerate, Nonresponse Followup), field workers are provided the opportunity to add newly discovered addresses and delete addresses that they are unable to find on the ground or that are represented multiple times in their workload.

All of these activities share the goal of having every housing unit included once and only once in the MAF, with its location properly reflected geographically.

Summary of our Current Knowledge of the Master Address File for 2010

At the beginning of my presentation I noted that there are three principal criteria for the evaluation of the Master Address File. We have indirect indicators of each of these, only in a preliminary fashion at the moment.

Coverage of Housing Units in the Country

- Relative to 2000, fewer governments participated in the Local Update of Census Addresses program to provide local updates of addresses; however, they did represent 92 percent of all addresses in the country. The comparison with 2000 is complicated by state government participation that included lower-level governments in 2010.
- The state and local governments provided addresses that form about 1.7 percent of the total valid addresses on the file after Address Canvassing. In addition, updates from local governments provided block codes for approximately 4 million addresses that were on the MAF but without a geocode.
- After Address Canvassing, the total number of units on the file is comparable to independent estimates of the housing unit count.

Address Completeness of Units on the File

Address Canvassing found about 2 million “other living quarters,” now being revisited to aid in the identification of individual living units within them. We expect many of these to revert to a single housing unit definition after the Group Quarters Validation step this fall.

We continue to evaluate the current status of the Master Address File, and I hope that I might present the final pre-Census evaluation to this committee in the future. Please note though that much work remains to ensure a complete address list for the 2010 Census. The New Construction Program, the Count Review Program, additional updates from the U.S. Postal Service, and our own field work during the enumeration will provide more updates (both adds and deletes) to the 2010 Census address list.

Future Activities and Risks

Over the next two months, hundreds of important tasks must be completed across all components of the decennial census program.

There are a number of external events that could lead to delays or operational problems, such as a major hurricane, a widespread outbreak of H1N1 flu, or a major, last-minute design change imposed upon the program.

Internally, some of the major activities and risks over the next 60 days include:

- Completing the Group Quarters Validation operation this month so that we can update our control files for the enumeration of these places next spring. The field operation is going well, and we believe our systems are ready to capture and process the results, but at this stage of the census we are on a tight schedule that must be adhered to.
- Opening 344 additional Local Census Offices (LCOs) by the end of December in order to be ready to implement the major operations of the census next spring. Although this effort is going well at the moment, at any location, we risk running into space build-out issues, equipment deployment issues, telecom issues, or even bankruptcy issues with the leaser.
- Beginning recruiting for our major field operations next year, including deployment of toll-free jobs lines to the LCOs. Although it appears that current economic conditions will make recruiting of qualified applicants easier than in past censuses, between now and next spring we must fill over 1 million temporary positions. This is a massive challenge, which can be affected by things outside our control, such as a change in the economy. We also know there are some places in the country with high employment rates where we may have more difficulty in attracting enough applicants.
- Continuing the work on Paper-Based Operations Control System (PBOCS), which has an aggressive system development lifecycle due to the applications de-scoped from Field Data Collection Automation (FDCA), and a fixed schedule prior to 2010 Census operations. We believe this is being managed well, and that we are on schedule for the deployment and use of these systems, but the schedule is very tight, with little room for any slippage.
- Preparing for and beginning production of the address label files for 2010 Census questionnaires and Advance Letters. Again, we believe this is on schedule for timely completion, but any significant problems with these efforts could jeopardize many aspects of the program.
- Completing a number of activities related to enhancements to our language outreach efforts. We believe these activities can be completed and integrated on schedule, but these efforts are on a tight schedule as well.

- Completing the first Operations Test and Dry Run for our Data Capture Centers and Call Centers. Developmental work and testing is going well, but we must stay on track to ensure these centers are ready next spring to capture and process census forms, and respond to public questions and assistance requests.

Conclusion

Mr. Chairman, the Census Bureau is very sensitive to the concerns you have expressed in the past with respect to what we are doing to address the undercount, to ensure we don't miss people in the upcoming Census. Continual maintenance and update of the Master Address File is one very important and extensive way we do that; we want to make sure every household that exists and is occupied receives a form to complete and mail back. Additionally, we want to work with our state, local and tribal government officials in a way that makes it easier, not harder, for them to share their information on new and existing addresses so we can get this right and not miss anyone. We think our process will allow us to do just that.

I thank the committee for this opportunity to testify and would be happy to answer any of your questions.