

*U.S. DEPARTMENT OF COMMERCE  
Office of Inspector General*



*U.S. Census Bureau*

*Census 2010: Dress Rehearsal of  
Address Canvassing Revealed  
Persistent Deficiencies in  
Approach to Updating the  
Master Address File*

*Final Report No. OSE-18599  
October 2008*

**FOR PUBLIC RELEASE**

*Office of Audit and Evaluation*





10/17/2008

**MEMORANDUM FOR:** Steve Murdock  
Director  
Census Bureau

*Judith J. Gordon*

**FROM:** Judith J. Gordon  
Assistant Inspector General  
For Audit and Evaluation

**SUBJECT:** *Census 2010: Dress Rehearsal of Address Canvassing  
Revealed Persistent Deficiencies in Approach to Updating  
the Master Address File*  
Final Report No. OSE-18599

This is our final report on our review of the 2008 dress rehearsal test of the bureau's planned address canvassing operation and the operation's impact on improving the master address file (MAF).

Our review revealed that (1) millions of ungeocoded Postal Service addresses in the MAF set the stage for missed housing units and increased workloads; (2) address canvassing cannot be relied on to adequately update, delete, or add addresses for certain types of housing units or residential communities, such as apartments and trailer parks; and (3) the bureau's modified quality control procedures may compromise the results of the quality control operation.

As you know, the actual decennial canvassing operation, which is scheduled for 2009, is the bureau's last opportunity to significantly improve the MAF. Our recommendations for improving the list take into account this limited time frame. But looking ahead to the 2020 census, we urge the bureau to replace this seriously limited operation with a decade-long, continuous improvement program.

We appreciate your response to the draft version of this report, and ask that you provide an action plan within 60 calendar days that details how you will implement our recommendations. We have addressed your response in summary at the end of this report and have attached it in full as appendix B.

We appreciate the cooperation and courtesies your staff at headquarters and in the field extended to us during our review. If you would like to discuss this report or the action plan, please feel free to call me at (202) 482-2754.

cc: Cynthia A. Glassman, Under Secretary for Economic Affairs  
Arnold Jackson, Associate Director for Decennial Census,  
U.S. Census Bureau  
Marilia A. Matos, Associate Director for Field Operations,  
U.S. Census Bureau  
Otto J. Wolff, Chief Financial Officer and Assistant Secretary for  
Administration, U.S. Department of Commerce  
Phyllis Van Tassel, Census Audit Liaison

## INTRODUCTION

To conduct the 2010 census, the Census Bureau will have to contact, via mail or in person, an estimated 134 million housing units, and will rely on its master address file (MAF) to do so. The MAF is intended to be a current, comprehensive list of every address in the nation—whether occupied or vacant. It is a compilation of, among other things, the bureau’s address list from Census 2000, U.S. Postal Service data, and local government address information provided primarily through an operation known as LUCA—Local Update of Census Addresses.

The bureau’s mechanism for collecting and tabulating decennial census data is to link (geocode) MAF addresses to the bureau’s digital map known as TIGER (Topologically Integrated Geographic Encoding and Referencing). A subset of the MAF is used to create the list of addresses for living quarters to be included in decennial operations and ultimately to count people where they live. Only geocoded addresses are sent to the address canvassing operation, and only geocoded addresses receive census questionnaires. Geocoding is how the bureau fulfills its mandate to count people in the right location for purposes such as redistricting.

The bureau conducts address canvassing to ensure the address list for the census is as accurate as possible. Estimated to cost more than \$500 million, this operation entails verifying, updating, or deleting addresses; adding missing addresses; updating streets on the TIGER maps; and geocoding every structure by assigning Global Positioning System (GPS) coordinates. During this massive operation, thousands of temporary decennial staff using handheld computers containing MAF addresses and TIGER maps try to locate every place a person could live or stay and ensure it is correctly recorded and geocoded.

Address canvassing for the upcoming decennial is scheduled for 2009. This is the bureau’s final major operation for improving the MAF prior to conducting the decennial.

***An accurate list is fundamental to an accurate count.*** The bureau describes “an accurate, comprehensive, and timely [address] list” as “one of the best predictors of a successful census.” If the list is incomplete or inaccurate, people may be missed or counted more than once. Errors in the MAF can also increase costs and the public burden by requiring visits to nonexistent or duplicate locations in a subsequent census operation known as nonresponse follow-up—the most costly operation in the decennial census.

Duplicate addresses occurred so frequently in Census 2000 that the bureau conducted an unplanned operation late in the census to identify them. It found that 6 million people may have been counted twice and it dropped 3.6 million of the suspected duplicate addresses from the census count. The National Academy of



Sciences cited post-decennial bureau studies that found a significant number of people had been dropped erroneously, while others who should have been eliminated were not.<sup>1</sup>

To reduce such problems in 2010, the bureau has been conducting ongoing MAF maintenance activities, including twice-yearly updates with Postal Service addresses. It had planned to conduct a number of other MAF improvement operations throughout the decade with the intent of reducing the scope of address canvassing. Census reported that budget cuts prevented it from conducting the research required to target areas for address canvassing prior to the decennial and from performing all of its planned address list improvement activities. So as in Census 2000, the bureau is depending on 100 percent address canvassing to correct the MAF before it begins the 2010 population count.

We had evaluated activities to update the MAF and related maps during the 2006 census test of planned decennial operations and identified missing, duplicate, and erroneous addresses in the list. During an actual census, these problems would hinder the bureau's ability to get a complete and accurate count. (Visit our [Census Reading Room](#) to view our related reports.)

We assessed these activities again in the 2008 dress rehearsal—reviewing information for 18,694 addresses from 125 assignment areas (100 randomly selected by Census and another 25 that we observed during dress rehearsal address canvassing). We evaluated the quality of the MAF for these areas before and after address canvassing and found that many of the issues we noted in 2006 remain. Although we cannot project an error rate to the overall MAF based on our sample, the regularity with which these problems surfaced suggests they may be significant and could reduce the accuracy of the count in 2010. Census believes other follow-up operations will resolve some of these errors, but none will provide a comprehensive review of—or fix for—the addresses to be used in the decennial. (See appendix A for additional details about our objective.)

## FINDINGS AND RECOMMENDATIONS

### I. Ungeocoded Addresses and Incorrect Zip Codes in the MAF Set the Stage for Problems in Address Canvassing

**Ungeocoded addresses.** According to Census officials, the MAF contains some 6 million ungeocoded addresses. If an address is not geocoded, it is not on the list given to address canvassers. The burden is then on canvassers to find and correctly record the information for these addresses—a burden the canceled MAF-

---

<sup>1</sup> *The 2000 Census: Counting Under Adversity*, Constance F. Citro, Daniel L. Cork, and Janet L. Norwood (eds.), Panel to Review the 2000 Census, National Research Council, The National Academies Press, 2004.

improvement operations were intended to reduce. Census will not mail questionnaires to any ungeocoded addresses that remain in the MAF after canvassing, even though the Postal Service can deliver most of them. And nonresponse follow-up enumerators only contact households that received but did not return a questionnaire. Therefore, residents at these ungeocoded addresses are unlikely to be counted.

**Outdated ZIP codes.** Several hundred addresses in our sample were duplicates except for the ZIP code. Up-to-date addresses were added by a LUCA operation conducted as part of dress rehearsal testing. But the updates were not recognized as the same housing unit in the MAF because the ZIP codes were different, so the LUCA addresses did not override the outdated entries. We brought this to the attention of Census staff, who told us that for 2010, they will change the way LUCA data is matched to the MAF with the objective of preventing duplicates.

## **II. Address Canvassing Remains an Unreliable Approach for Improving the MAF**

As in our earlier reviews, we found that address canvassing cannot be counted on to reliably record the information needed to geocode the millions of ungeocoded addresses in the MAF, identify and add missing units, or eliminate certain address errors. The operation is especially challenged when dealing with (1) nontraditional mailing addresses—for example, multiunit residences whose individual household addresses do not conform to the typical street number and name format; (2) “hidden” dwellings, such as sheds and makeshift garage apartments; and (3) trailer parks that display both lot numbers corresponding to a unit’s physical space within the park and unit numbers that are part of the mailing address. The following examples from dress rehearsal canvassing illustrate some of the problems we observed:

- *Entire apartment complex omitted from canvassing address list.* A 352-unit apartment complex in our sample that was built after 2000 was in the MAF but not geocoded and was therefore not on the address list given to the employee assigned to canvass the area. The canvasser correctly added 346 addresses, but mistakenly missed 6 apartment units and added 1 that did not exist. The 6 missed apartments therefore remain ungeocoded. If these omissions are not corrected during address canvassing for 2010, the units will not be mailed a census questionnaire, increasing the risk that their residents will not be counted. Likewise, the nonexistent unit will be added to the nonresponse follow-up workload. Occupants of housing units missed in the address canvassing operation for 2010 face an increased risk of not being counted.

- *Certain ungeocoded mailing addresses not corrected in dress rehearsal will likely remain ungeocoded in 2010.* The outcome of canvassing a 391-unit retirement community in our sample illustrates two possible problems that could diminish the accuracy of the 2010 count. Canvassing lists for the retirement community contained “location” addresses—street names and unit numbers (e.g., 3629 E. Adams Avenue)—which in this case do not correspond to the Postal Service mailing addresses. Instead the Postal Service uses *the complex name* and unit number, known as the “vanity” address (e.g., 3629 ABC Apartments).<sup>2</sup> The vanity addresses exist in the MAF (as provided by Postal Service updates) but are not geocoded, or linked, to a specific block location, so they do not appear on canvassers’ lists.

Unit numbers in this complex were displayed on each door. During dress rehearsal, some canvassers knocked on doors as instructed by Census’s address canvassing procedures to verify address information and learned that the mailing address contained the complex name rather than the street. They added this information for some units—now allowing the Postal Service addresses to geocode—but not for others. The first problem we note is that the remaining units may not receive Census questionnaires in 2010 if the Postal Service does not recognize the location address as a valid mailing address.

The second problem is potentially more significant: we learned that canvassers for 2010 will collect mailing addresses only when housing unit numbers are not clearly displayed, which means that none of the addresses for complexes like the one above—that have visible unit numbers but use vanity addresses for mailing purposes—will be corrected. In addition, even if a canvasser should happen to learn that complexes with displayed unit numbers use vanity mailing addresses while the MAF contains only the location addresses, the bureau has made a software change to the handhelds that will prevent the canvasser from making the correction. Consequently, these complexes may not receive a questionnaire unless the bureau finds a way to identify ungeocoded valid Postal Service mailing addresses that exist in the MAF and match them to their corresponding house number and street name addresses. No one knows how many complexes/households fall into this category and therefore how many people might not receive questionnaires, but the impact will be an increased workload for the costly nonresponse follow-up operation.

- *Insufficient address information for “hidden” dwellings.* Address canvassing is the primary means for identifying hidden living quarters. Yet we found that

---

<sup>2</sup> For most households, the location address is the mailing address. But some multiunit dwellings use so-called vanity mailing addresses—identifying the individual units by the complex name rather than the street on which the building resides. When the vanity address is the mailing address, the postal service will sometimes deliver mail to the location address, but cannot always be counted on to do so.

almost 600 such units identified by dress rehearsal canvassers were not added to the master address file. Because canvassers did not sufficiently differentiate the hidden dwellings from the primary residence at the same location by placing a unique alpha or numeric identifier in the handheld’s “unit” field, the hidden units were considered duplicate addresses. Canvassers had correctly added location descriptions to the handhelds, such as “apartment above garage” or “trailer in the side yard.” But they left the unit field blank, so the record was rejected. The handheld contains no prompt to make it clear that the unit field must be completed for these types of dwellings; and clear instructions—both in training sessions and instructional manuals—are lacking. We briefed Census officials about this finding, and they are working with the handheld computer contractor to resolve the problem.

- *Trailer park lot numbers incorrectly recorded as unit numbers.* We found numerous trailer parks for which dress rehearsal canvassing lists contained duplicate addresses as a result of a Census 2000 list error. Even though each trailer had a unique house number and street name address, canvassers for the 2000 decennial had incorrectly added lot numbers—which are not part of the mailing address—as “unit” numbers. Subsequent Postal Service and LUCA updates added new MAF records with the accurate addresses for these dwellings, but did not override the erroneous entries generated in the 2000 canvassing operation. Because the procedure for resolving duplicates was unclear, 2008 canvassers typically kept the erroneous address and deleted the Postal Service one. We expect the same outcome from address canvassing for 2010, and future MAF update operations will likely add the correct address information again. To break this cycle, canvassers must understand that the unit field is completed only for dwellings that share the *same house number* and street name address, but have unique identifiers for their individual units. All other location details belong in the location description field (figure 1).



Figure 1. Handheld computer screen for adding an address

### III. Modifications to the Quality Control Operation Have Not Been Thoroughly Evaluated

**Modified procedures may compromise the operation’s results.** Census has tested and implemented a number of quality control improvements since 2000, most important, making quality control independent of production. However, many aspects of the quality control operation were new for the dress rehearsal, and automation problems prevented the bureau from collecting sufficient data to fully evaluate them. Nevertheless, these changes will be used in 2010.



One key change is the procedure for handling assignment areas that fail the quality check. In 2000, if an assignment area failed quality control, it was returned to production staff to be re-canvassed. For 2010, quality control staff does the check *and* the re-canvassing. These dual responsibilities may compromise the operation's results: quality control staff could choose to shortcut the re-canvassing process and therefore fail to make needed corrections or could identify nonexistent errors in order to increase their work hours. This is not just a hypothetical risk; we observed an instance of the former problem during dress rehearsal. Census should closely monitor quality control pass/fail rates for assignment areas nationwide—taking note of unusually high or low pass/fail results.

***Quality control staff introduced errors.*** We found examples of quality control staff reinstating addresses that had been correctly deleted by the canvasser and incorrectly adding addresses to canvassers' lists. In one example, 17 addresses correctly verified by a canvasser in one block were added to an adjacent block during quality control. So two sets of identical addresses now exist on the same street. In 2000, geography matching operations in headquarters would have attempted to recognize and eliminate the identical addresses from the second block. But for 2010, Census has decided it must keep these duplicate addresses as valid housing units to avoid missing a residence. The bureau informed us that if it receives completed questionnaires from both, the addresses will not be identified as duplicates since the housing units are in separate blocks. Such instances will result in overcounting.

Bureau officials contend that once canvassing is over, they cannot determine which address is in the correct block. We disagree. For the example above, we used county geographic information system sites to confirm that the block as initially verified by the address canvasser was correct. We also determined that the GPS coordinates for the addresses added by quality control were in the middle of the road, not at the physical location of actual dwellings. The bureau intends to track the frequency of these errors during address canvassing for the 2010 census and will revisit the issue, if warranted.

### **Recommendations**

To improve the master address file for 2010, Census should do the following:

1. Mail questionnaires to ungeocoded postal addresses, and geocode as many addresses as possible for forms returned from valid housing units.
2. Explore ways to identify ungeocoded vanity and other unique mailing addresses in the MAF, and link them to corresponding location addresses.
3. Enhance instructions and training to advise canvassers that
  - a. the "unit" field must be completed for every individual dwelling within a multiunit residence—including hidden dwellings, and

- b. secondary identifying details—such as lot numbers in trailer parks—must be entered in the “location description” field.
4. Closely monitor the level of quality control pass/fail rates nationwide to detect quality control staff that may be improperly passing or recanvassing assignment areas.
5. Reexamine the decision to keep identical MAF addresses that occur in different blocks within the same street.

These recommendations are tailored to account for the decennial scheduling and automation problems facing the bureau in preparing for the 2010 census. To eliminate the recurring problems we have consistently identified and improve the MAF for 2020, we further recommend that Census regularly update the file throughout the decade in lieu of 100 percent address canvassing.

### **Bureau Response and OIG Comments**

With respect to our first recommendation, Census officials stated that, while they share our concern about ungeocoded addresses in the MAF, mailing to these addresses increases the likelihood of duplication, which increases the risk of an overcount. But we note that their current plan poses the risk for undercounting. We believe the bureau can balance both risks with targeted mailings to specific groups of ungeocoded addresses, such as those added after 2000, or to areas that are likely to have a large number of these addresses, such as communities that experienced rapid growth and were not covered by LUCA.

Officials agreed with the merit of recommendations 2 and 5, but stated that time and resource constraints precluded them from conducting additional research.

Census concurred with recommendations 3(a) and 4. It did not concur with 3(b), but we believe the bureau misinterpreted our example, so we have clarified our meaning in the final draft and ask that the bureau address this recommendation in its action plan.

Finally, the bureau agreed with our recommendation for the 2020 decennial—that it regularly update the MAF throughout the decade in lieu of 100 percent address canvassing—but said doing would so hinge on whether it obtains the necessary funding. We urge Census to make this a top priority and take action early in the decade to secure the needed funding.

## **APPENDIX A: OBJECTIVES, SCOPE, AND METHODOLOGY**

We originally intended to assess how well address canvassers followed procedures during the 2008 dress rehearsal and made corrections on the handheld computers. We also planned to assess whether the operation's redesigned quality control component was effective. But problems with the handhelds forced Census to modify its objectives for the dress rehearsal, and we in turn had to modify ours, focusing primarily on the quality of the list. Because Census did not agree to provide assignment area data to us electronically, we could not conduct various sorts, searches, and other automated analyses. Consequently our findings are based on a manual review of 1,000 printed pages containing nearly 19,000 addresses from the North Carolina dress rehearsal site. We have raised concerns to the Department and the Census Bureau about restrictions placed on OIG's access to Title 13 information and are awaiting a final response from Census regarding our concerns.

We conducted this review from April 2007 to July 2008 in accordance with the *Quality Standards for Inspections* (rev. January 2005) issued by the President's Council on Integrity and Efficiency, and under the authority of the IG Act of 1978, as amended, and Department Organization Order 10-13 (dated August 31, 2006).

APPENDIX B: CENSUS BUREAU'S RESPONSE



UNITED STATES DEPARTMENT OF COMMERCE  
Economics and Statistics Administration  
U.S. Census Bureau  
Washington, DC 20233-0001  
OFFICE OF THE DIRECTOR

OCT 01 2008

MEMORANDUM FOR Judith J. Gordon  
Assistant Inspector General for Audit and Evaluation

Through: Cynthia A. Glassman *CAG*  
Under Secretary for Economic Affairs

From: Steve H. Murdock *Steve H. Murdock*  
Director

Subject: Comments on Recommendations Within the Office of Inspector  
General's Second Draft Report Entitled *2010 Decennial Census:  
Dress Rehearsal of Address Canvassing Revealed Persistent  
Deficiencies in Approach to Updating the Master Address File  
(OSE-18559)*

The attached is in response to your request of October 1, 2008, for specific comments on the recommendations contained in the above-referenced report. We appreciate the opportunity to review the report prior to publication.

Attachment

cc: US/EA

U.S. Census Bureau Comments on the  
Second Draft Report No. OSE-18559/September 2008  
*2010 Decennial Census: Dress Rehearsal of Address Canvassing Revealed Persistent  
Deficiencies in Approach to Updating the Master Address File*  
October 7, 2008

**General Comment**

The U.S. Census Bureau concurs with the Office of Inspector General (OIG) as to the relative value of continuous updating and error correction as opposed to the end-of-the-decade national canvassing approach. We would welcome the OIG support for defending the budget necessary to implement a rigorous continual Master Address File updating program, supported by a strong research program next decade. Further, the Census Bureau shares the OIG concerns about duplicates, misses, and ungeocoded housing units and some of the operational risks facing us in Address Canvassing. Consequently, we are in the midst of developing and implementing a comprehensive Address Canvassing testing program, including an operational field test, as well as completing mitigation and contingency plans to address risks to the success of Address Canvassing and to census data quality. As testing and risk management activities unfold, we would welcome the OIG's input and support. We appreciate the recommendations you have made and would like to offer the following comments.

**OIG Recommendations**

*To improve the Master Address File (MAF) for 2010, Census should do the following:*

**Recommendation 1. Mail questionnaires to ungeocoded postal addresses, and geocode as many addresses as possible for forms returned from valid housing units.**

**Census Bureau Comment:** We recognize and share the concern regarding the ungeocoded addresses. We are examining options as contingencies for the event that a significant number of ungeocoded addresses remain after the completion of Address Canvassing. Your recommendation is being considered; however, mailing questionnaires to ungeocoded addresses will not result in the geocoding of those addresses. In fact, mailing to all ungeocoded addresses increases the likelihood of duplication in the census because we may not be able to determine for certain whether the addresses have already been added by Address Canvassing in a different form. Using the example in the report, if Address Canvassing added 3629 East Adams Avenue, we would create duplicate records by adding ungeocoded postal addresses for 3629 ABC Apartments.

**Recommendation 2. Explore ways to identify ungeocoded vanity and other unique mailing addresses in the MAF, and link them to corresponding location addresses.**



**Census Bureau Comment:** We concur that this issue warrants research. However, given the very limited time that remains prior to the start of the 2010 Census Address Canvassing operation and the focus of our resources on preparatory activities, this research is not feasible at this time.

**Recommendation 3. Enhance instructions and training to advise canvassers that—**

**a. the “unit” field must be completed for every individual dwelling within a multiunit residence—including hidden dwellings, and**

**Census Bureau Comment:** We concur with this recommendation, but with a caveat. A blank unit designation is considered unique from other non-blank unit designations. Therefore, there could be up to one blank unit designation per multiunit.

**b. secondary identifying details—such as lot numbers in trailer parks—must be entered in the “location description” field.**

**Census Bureau Comment:** We do not agree with this recommendation. There will be instances when identifying details, such as lot numbers in trailer parks, are necessary to uniquely identify each address. Instructing listers that this information must be entered in the “location description” field, rather than in the “unit designation” field, will result in duplicate addresses during the MAF update process.

We have uniqueness requirements in the Address Canvassing hand-held software. We require something different for every unit with the same basic street address, where “blank” is one of the possibilities. In addition, if a mobile home park has the same basic street address, then the software will require something different in each unit descriptor. Addresses that are not entered with this amount of information will fail both our software uniqueness requirement and our MAF update uniqueness requirement.

**Recommendation 4. Closely monitor the level of quality control pass/fail rates nationwide to detect quality control staff who may be improperly passing or recanvassing assignment areas.**

**Census Bureau Comment:** We concur with this recommendation.

**Recommendation 5. Reexamine the decision to keep identical MAF addresses that occur in different blocks within the same street.**

**Census Bureau Comment:** Based on your recommendation, in July 2008, the Address List Consistency Group re-examined our decision to keep identical MAF addresses that occur in different blocks. To determine which block the address is located in would require additional research. We do not have the resources necessary to implement a national scale research effort within the time frame available between the completion of

Address Canvassing (July 2009) to the start of Group Quarters Validation (September 2009) and the production of the initial mail file. As a result, we decided to continue with our current approach due to concerns about under-coverage. However, we do plan to track the level of occurrence during Address Canvassing and, if warranted, will revisit this issue again.

*To eliminate the recurring problems we have consistently identified and improve the MAF for 2020, we further recommend that:*

**Recommendation 6. Census [should] regularly update the file throughout the decade in lieu of 100 percent address canvassing.**

**Census Bureau Comment:** We concur with this recommendation. If funding will be available, we will conduct the research necessary to determine if we could reduce the scope of Address Canvassing and improve the address list throughout the decade leading up to the 2020 Census.