



U.S. CENSUS BUREAU

High-Quality Maps and Accurate Addresses Are Needed to Achieve Census 2020 Cost-Saving Goals

FINAL REPORT NO. OIG-12-024-I

MAY 10, 2012

U.S. Department of Commerce
Office of Inspector General
Office of Audit and Evaluation

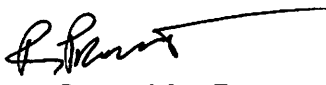
For Public Release





May 10, 2012

MEMORANDUM FOR: Robert M. Groves
Director, U.S. Census Bureau

FROM: Ron Prevost 
Assistant Inspector General for Economic
and Statistical Program Assessment

SUBJECT: High-Quality Maps and Accurate Addresses Are Needed to
Achieve Census 2020 Cost-Saving Goals
Final Report No. OIG-12-024-1

We are providing our final report on the effectiveness of the Census Bureau's MAF/TIGER database (MTdb). Our objectives were to (1) review the bureau's progress toward achieving the objectives of its \$496 million 2010 geographic initiative (2010 MAF/TIGER Enhancement Program), (2) evaluate the impact of various address-updating operations on the address file and map database and identify trends that introduced error, and (3) review Census's procedures for updating the map and address files.

We reviewed the bureau's progress on the MAF/TIGER Enhancement Program and found that two of the objectives had not been fully realized: developing a measure to assess MTdb quality at low levels of geography and enhancing geographic partner programs with tribal, state, county, and local governments. Two other areas that require improvement are the number of ungeocoded addresses (new addresses that are not linked to a location on a map) in the MTdb and the current address-updating process, which accepts the most recent address changes without adequate verification. To implement cost-saving goals for the 2020 census, the bureau must address these issues.

We have received your official response and technical comments to our draft report. Where appropriate, we have modified this final report based on the information provided to us. The official response is included as appendix B. (We summarized your response and OIG comments on page 12.) The final report will be posted on the OIG website pursuant to section 8L of the Inspector General Act of 1978, as amended.

In accordance with the Department Administrative Order 213-5, please provide us with your action plan within 60 days of the date of this memorandum. We thank you and your staff for the courtesies extended to us during this review. Please direct any inquiries regarding the report to me at (202) 482-3052 or Carol Rice at (202) 482-6020.

Attachment

cc: Mark E. Doms, Chief Economist, Economics and Statistics Administration
Frank Vitrano, Associate Director for the 2020 Census, U.S. Census Bureau
Pam Moulder, Senior Program Analyst, Economics and Statistics Administration
Adam Miller, Audit Liaison, U.S. Census Bureau



Report In Brief

MAY 10, 2012

Background

The Census Bureau maintains a database containing a complete list of all living quarters (the master address file, or MAF) and geospatial data (the topologically integrated geographic encoding and referencing system, or TIGER) of the nation to use in all demographic and decennial programs. The bureau's method of collecting and tabulating decennial census data is to link (geographically encode, or geocode) MAF addresses to TIGER.

The Census budget for maintaining the MAF/TIGER database (MTdb) was \$425 million for the period leading up to the decennial census (fiscal years 2000 through 2010). Census spent another \$496 million on a 2010 geographic initiative: an 8-year effort ending in 2010 to improve the accuracy of the MTdb. Finally, in 2009, a \$444 million operation to visit and verify or add every place a person lives or could live was conducted. Totalling nearly \$1.4 billion, these combined efforts produced the 2010 decennial census address list.

For the 2020 census, the bureau intends to continuously update the MTdb, which will improve the address lists and maps throughout the decade and support a less costly targeted address-canvassing operation. Toward this effort, the bureau has introduced a \$407 million 2020 geographic initiative.

Why We Did This Review

Our objectives were to (1) review the bureau's progress on its 2010 geographic initiative, (2) evaluate the impact of various address-updating operations on the MTdb and identify trends that introduced error, and (3) review Census's procedures for updating the map and address files.

U.S. CENSUS BUREAU

High-Quality Maps and Accurate Addresses Are Needed to Achieve Census 2020 Cost-Saving Goals

OIG-12-024-I

WHAT WE FOUND

The 2010 geographic initiative's efforts to produce an adequate measure to assess MTdb quality were unsuccessful. In addition, the program's goal of updating address and map information from tribal, state, county, and local government partners was not fully realized. Both of these goals must be met to implement a 2020 decennial census address-canvassing operation with reduced costs. Furthermore, 3.5 million ungeocoded records existed in the MAF as of June 2011, and that number is likely to rise, as it did during Census 2010. Without maintenance of the MTdb by continuous geocoding throughout the decade, the bureau will again have to rely on an expensive end-of-decade operation. Finally, the MAF updating process of accepting more recent address changes without adequate verification may result in a lower quality address list.

WHAT WE RECOMMEND

1. Develop an MTdb measure for determining address list quality at a low level of geography that (a) provides a fair and equal opportunity for targeting selection, (b) drives selection and planning decisions, and (c) is well-documented and verifiable.
2. Work with the Department to determine the feasibility of improving methods of sharing MTdb information throughout the decade with governmental entities (partners) to create a uniform, national address list.
3. Investigate and remedy the exclusion of 500,000 ungeocoded address records, which had been designated as valid U.S. Postal Service delivery addresses, from the 2010 census.
4. Conduct the necessary research, develop a proven methodology, and allocate the necessary funds to continuously reduce the number of ungeocoded records throughout the decade.
5. Develop and implement quality indicator tools, including use of administrative records, to ensure that updates to the MAF are accurate.

Contents

Introduction 1

Findings and Recommendations 3

 I. The 2010 Census Geographic Initiative Fell Short, Raising Concerns About Cost-Saving Measures for the 2020 Census 3

 A. The measure to assess MTdb quality was not developed but is needed to reengineer the next decennial..... 3

 B. Strained partner relations and Title 13 restrictions could hamper efforts to maintain an updated, accurate address list throughout the decade 5

 II. Census Relied on Expensive End-of-Decade Operations to Validate Millions of Postal Service Addresses in the MTdb 7

 III. Census’s Process of Updating Addresses by Using the Most Recent Address May Result in a Lower Quality Address List..... 8

Recommendations 11

Summary of Agency and OIG Comments..... 12

Appendix A: Objectives, Scope, and Methodology 13

Appendix B: Agency Response 14

*COVER: Detail of fisheries pediment,
U.S. Department of Commerce headquarters,
by sculptor James Earle Fraser, 1934*

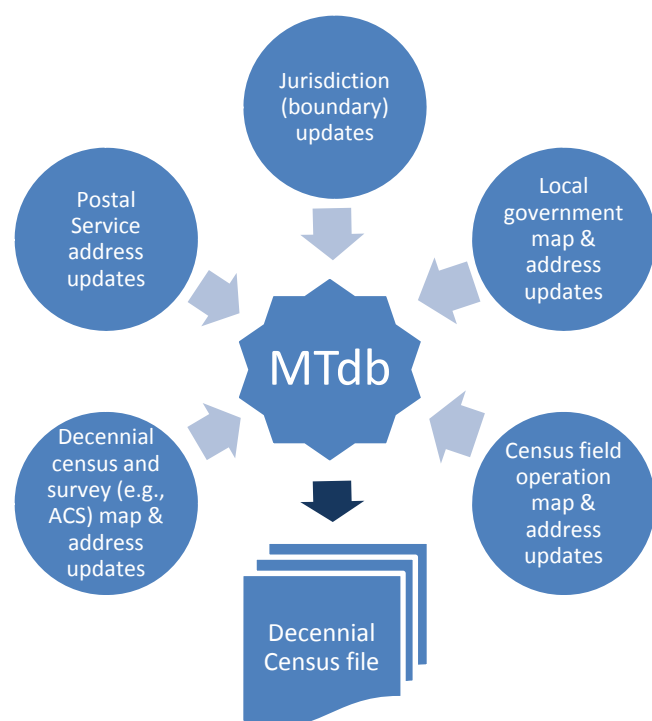
Introduction

As the federal government’s largest statistical agency, the U.S. Census Bureau manages a full range of demographic programs—including the decennial census, the American Community Survey (ACS), and other population and income surveys—that drive major statistical programs across multiple federal agencies. Results of its surveys and programs serve as resources for determining how more than \$400 billion of federal funds are distributed annually and how seats in the U.S. House of Representatives are distributed every 10 years and for a broad range of uses by businesses and the public.

Essential to executing this work, the bureau maintains a database that contains a complete list of all living quarters (referred to as the master address file, or MAF) and geospatial data¹ (referred to as the topologically integrated geographic encoding and referencing system, or TIGER) of the nation to use in all Census Bureau demographic and decennial programs. The bureau’s method for collecting and tabulating decennial census data is to link (geographically encode, or geocode) MAF addresses to TIGER.

The MAF/TIGER database (MTdb) is the backbone of Census demographic operations and is critical to implementing bureau surveys and generating data products (i.e., Census’s demographic results are collected at the household level, which allows data users to answer any

Figure I. Multiple Sources Update the MTdb



number of questions, such as the number of school-age children in the nation, in a particular city, or even in a particular neighborhood). Most demographic surveys, including the decennial census, have unique requirements that determine the list of addresses extracted from the MTdb.

Figure I illustrates how multiple sources update the MTdb to produce the lists of addresses used to collect household information. Because there is no single source for updating data in the MTdb, the bureau must coordinate with the providers of multiple data sources (referred to as partners), most of which are outside of Census’s direct control, or conduct its own operation to verify and update addresses and maps. Since each partner collects data for a different purpose and has different quality control practices, changes (e.g., revisions in housing unit addresses, roads

¹ The geospatial data are displayed graphically as maps. Therefore, throughout this report, we refer to the geospatial data as “maps.”

and highways, and government boundaries) come to the bureau with varying levels of precision.

The Census division charged with maintaining the MTdb had a \$425 million budget spread out across the period leading up to the decennial census (fiscal years 2000 through 2010). The division spent an additional \$496 million on a 2010 geographic initiative (MAF/TIGER Enhancement Program), an 8-year effort ending in 2010² that sought to improve the accuracy of the MTdb. Finally, in 2009, the bureau initiated a \$444 million operation to visit and verify or add every place a person lives or could live. Totalling nearly \$1.4 billion, these combined efforts produced the 2010 decennial census address list.

To reduce costs for the 2020 census, the bureau intends to continuously implement a more robust update of the MTdb. A continuously updated, accurate MTdb would improve the address lists and maps throughout the decade and support a less costly targeted address-canvassing operation. To support this effort, the bureau has introduced a \$407 million 2020 geographic initiative (Geographic Support System).

This report presents the results of our evaluation of the Census Bureau's efforts to maintain updated, accurate addresses and maps of all living quarters.³ Our objectives were to

1. review the bureau's progress toward achieving the objectives of its \$496 million 2010 geographic initiative,
2. evaluate the impact of various address-updating operations on the MTdb and identify trends that introduced error, and
3. review Census's procedures for updating the map and address files.

The purpose of our evaluation is to inform the next decennial. Additional information on our scope and methodology can be found in appendix A.

² Some limited funding for one of the objectives was planned through FY 2012.

³ The bureau defines living quarters as "any site where people live, stay, or could live."

Findings and Recommendations

I. The 2010 Census Geographic Initiative Fell Short, Raising Concerns About Cost-Saving Measures for the 2020 Census

Of the nearly \$1.4 billion spent on Census geographic programs over the past decade, \$496 million was for the 2010 geographic initiative. The main objective of this program was to align all street features to GPS map coordinates—a prerequisite for collecting coordinate locations of residential structures during the address-canvassing operation. By aligning streets and residential structures to GPS, the bureau sought to improve the accuracy of the coordinate information in the MTdb. By 2008, the streets in every county in the United States were aligned to GPS map coordinates. The bureau accomplished two other objectives: integrating the MAF and TIGER databases in a single Oracle database and developing and implementing, through FY 2006, a field operation to capture addresses in predominantly rural areas of the nation where city-style addresses generally are not used for mail delivery (referred to as the Community Address Updating System).⁴ However, the remaining two objectives—improving MAF and TIGER quality metrics and developing and enhancing geographic partnership (tribal, state, county, and local governments) programs throughout the decade—fell short.

Leading up to the next decennial census, the bureau received approval in FY 2011 for a 2020 geographic initiative with three overall goals (continual spatial feature updates, improved address coverage, and enhanced quality assessment and evaluation) for a total expected life-cycle cost of \$407 million. The initiative contained several new as well as some familiar 2010 objectives (see table I) that aim to address last decade's shortfalls.

A. The measure to assess MTdb quality was not developed but is needed to reengineer the next decennial

An objective of the 2010 geographic initiative was to implement a comprehensive plan for periodic MAF/TIGER evaluation that would identify areas where the address list was missing housing units or was outdated. However, the measure developed to achieve this goal (National Estimate of Coverage) evaluated MTdb quality at the state level. This \$44 million effort was not at a low enough level of geography to identify specific areas requiring improvement. To ensure a complete and accurate address list for 2010 decennial, the bureau implemented a \$444 million address-canvassing operation. This operation sent temporary census workers out to systematically traverse every street in the nation to add, delete, and correct addresses for places a person lives or could live and correct maps to ensure their completeness and accuracy.

⁴ In FY 2007, the Community Address Updating System was temporarily stopped because of funding constraints. The program resumed in FY 2010.

Table I. Census 2010 and 2020 Geographic Initiative Objectives and Status

2010 Geographic Initiative (\$496 million) MAF/TIGER Enhancement Program	2020 Geographic Initiative (\$407 million) Geographic Support System
Correct the street center line location for every county in the nation.	Continual spatial feature updates: add new or improve existing street attributes (e.g., street names and address ranges) and correct and improve spatially inaccurate streets.
<i>Objective met</i> —all counties were completed as planned.	<i>Ongoing.</i>
Create a modern processing environment by integrating the MAF and TIGER databases.	No corresponding goal.
<i>Objective met</i> —Commercial off-the-shelf software combined the two databases.	
Develop and enhance geographic partnership programs.	Improved address coverage: improve address list using updates from state, local, and tribal governments.
<i>Objective partially met</i> —a Web-based program for partners to upload data was deployed, but strained partner relations exist.	<i>Ongoing.</i>
Implement the Community Address Updating System Program.	Improved address coverage: update the address inventory not covered by the U.S. Postal Service.
<i>Objective met</i> —the Community Address Updating System Program was developed and implemented.	<i>Ongoing.</i>
Improve quality metrics.	Enhanced quality assessment and evaluation: assess address quality to target areas for update and refine the address universe; assess street location and attribute accuracy from various sources to guide updating decisions.
<i>Objective not met</i> —the measure developed was unable to assess MTdb quality.	<i>Ongoing.</i>
No corresponding goal.	Improved address coverage: assign locations to new addresses so they can be used in censuses and surveys, develop and implement methods to update the group quarter inventory of addresses, and update and maintain the Puerto Rico address list.
	<i>Ongoing.</i>
Source: OIG analysis of U.S. Census Bureau information	

A major reengineering effort for the 2020 census is to shift from a complete address-canvassing operation to a targeted operation. These efforts are in line with prior OIG recommendations: in *Census 2010: Final Report to Congress*⁵ we recommended that the bureau avoid a large-scale end-of-decade field operation by continuously updating the address lists and maps. A targeted operation of 5 to 20 percent of the addresses, according to the bureau, could result in a savings of between \$373 and \$442 million.⁶ These savings could be even greater—for the bureau and

⁵ U.S. Department of Commerce Office of Inspector General, June 27, 2011. *Census 2010: Final Report to Congress*, OIG-11-0030-I. Washington, D.C.: Department of Commerce OIG.

⁶ U.S. Census Bureau, *G.S.S. Initiative Cost Reduction* (PowerPoint presentation).

local governments—if an accurate address list also results in a smaller end-of-decade Local Update of Census Addresses (LUCA) program.

The decision to conduct a targeted operation hinges on establishment of measures that can determine address list quality at low levels of geography (such as the block or census tract⁷ level). Given the lack of success in developing a usable MTdb assessment measure last decade, the bureau has reinstated the objective in the 2020 geographic initiative. In addition to developing a measure, the bureau must ensure that stakeholders accept the measure and are confident that it accurately represents the existing housing inventory. Essentially, the bureau is determining which areas have an accurate inventory and require no additional fieldwork—a decision that could raise local, state, and congressional concerns about quality and fairness. We recommend that the bureau develop a measure that

- provides a fair and equal opportunity for an area to be included or excluded in a targeted address-canvassing operation,
- drives the selection and planning decisions for a targeted address-canvassing operation, and
- is well-documented and verifiable.

B. Strained partner relations and Title 13 restrictions could hamper efforts to maintain an updated, accurate address list throughout the decade

The 2010 census geographic initiative allocated \$43.7 million in funds last decade to develop and enhance geographic partnership programs. The bureau stated that the delivery of a Web-based application that managed partner data—allowing Census staff to gather information on partners, post communications via a secure site, and permit partners to download/upload data (such as boundary changes) to a secure server—satisfied this goal. Although the Web application was delivered on schedule, it appears that the primary goal, to develop and enhance geographic partnership programs, was not fully achieved. During interviews, staff from 6 of the 12 Census regional offices who work with the partners unanimously reported partner dissatisfaction, specifically citing the following:

- rigid schedule of Census requests,

LUCA is a legislatively mandated decennial census geographic partnership program that invites local, state, and tribal governments to review and correct Census's list of residential housing units and group quarters addresses, which will be used to deliver questionnaires in their jurisdiction.

⁷ Census tracts are small, relatively permanent statistical subdivisions of a county; boundaries normally follow visible features but may follow governmental unit boundaries. As of the 2010 census, the bureau tabulated 73,057 census tracts in the United States.

- lack of notice of upcoming projects, thereby not allowing partners to properly allocate resources (partners indicated they would like to know about Census requirements before the beginning of the fiscal year),
- incompatibility of Census and partner electronic data formats, and
- perception that Census does not fully use partner data.

Partners are often motivated to participate in the bureau's geography programs, because they realize the benefits of ensuring that their jurisdictions are correctly counted. For example, the number of children in low-income families is one basis for allocating federal funds to each school district, so it is imperative that this number is accurate. However, partner dissatisfaction, in addition to budget constraints, may lead to decreased participation and the loss of "free" updates (only in that there are no field data collection activities and cost) to the bureau.

Currently, the sharing of address lists between Census and tribal, state, county, and local governments, and even other federal government agencies, is limited. With a few very narrow exceptions, Title 13 forbids the Census from disclosing information furnished by respondents and imposes significant penalties—including criminal penalties—for violation of these confidentiality requirements.⁸ Consequently, although partners share addresses with Census, Census is restricted by law from reciprocating with many partners. The Census Address List Improvement Act of 1994⁹ did amend Title 13 to provide a limited exception to these restrictions for local governments. That act authorized the LUCA program, which allows two-way sharing of addresses information between the bureau and officials designated by local government units.¹⁰ However, the LUCA program has several requirements. The bureau must publish address standards, develop and publish a timeline for reviewing submissions, and provide a response back to participants. In addition, the Office of Management and Budget is required to implement an appeals process through which address-sharing disputes among the bureau and participants may be resolved. The 2010 LUCA program occurred from January 2007 through March 2010. Although it may require legislative action, a more informal method of two-way sharing of address lists earlier in the decade could improve address updating and geocoding as well as the cost-effectiveness of Census's demographic censuses and surveys throughout the decade.

The 2020 geographic initiative goal to use updates from state, local, and tribal governments supports yearly processing of partner data and creating a Web-based interface and new software application to accept and upload various partner formats. However, to move toward a continuously updated and more accurate MAF, we suggest that Census determine the feasibility of sharing MTdb information throughout the decade with governmental entities (partners) to create a uniform, national address list. If the 2020 census cost reduction strategy is to be realized, the bureau must ensure the MTdb is continually updated with partner data, acknowledging the risk that acquiring this data may be more challenging if partner resources diminish.

⁸ 13 U.S.C. § 9.

⁹ Pub. L. No. 103-430 (1994).

¹⁰ See 13 U.S.C. § 16.

II. Census Relied on Expensive End-of-Decade Operations to Validate Millions of Postal Service Addresses in the MTdb

Allocating population and housing to the correct location is the bureau's way of satisfying its constitutional mandate to count people where they live for purposes such as redistricting and apportionment. Geocoding is the process by which housing units are associated with a location on the ground. Ungeocoded records occur in the MTdb when a mailing address cannot be linked to a location on a map. Ungeocoded addresses are excluded from the decennial census and some bureau survey samples, increasing the likelihood that those housing units will not receive a questionnaire.

As shown in figure 2, the Census Bureau informed us that 10 million ungeocoded residential addresses existed in the MAF in March 2008. We tracked the volume of those ungeocoded records throughout the remainder of the decennial. At the end of the 2010 census operations, after all housing units were accounted for, 500,000 of the 10 million records identified in 2008 remained ungeocoded in the MTdb. At the same time, semi-annual updates from the Postal Service were adding more ungeocoded records to the MTdb. For some of these records, a valid, duplicate record may have existed in the MAF. For example, we previously found that Postal Service addresses may use a residential complex name rather than a house number/street name address (e.g., ABC Apartments instead of 10 Main Street).¹¹ In the MTdb, house number and street name addresses are required to electronically match an address to the map.

However, it is unlikely that 500,000 duplicate records exist. To determine other possible reasons for why these housing units were not accounted for in the MTdb, we compared the records against a number of demographic characteristics—such as population density, population and housing changes between 2000 and 2010, ethnic composition, vacant housing units, median income, and unemployment rate—but did not find any evidence that strongly predicted where or why ungeocoded records were likely to occur. We suggest that the bureau investigate the exclusion of these records, which had been designated as valid Postal Service delivery addresses, in the final decennial housing count. Identifying characteristics specific to these records would

Figure 2. Impact of End-of-Decade Activities on 10 Million Ungeocoded Records



¹¹ U.S. Department of Commerce Office of Inspector General, October 17, 2008. *Census 2010: Dress Rehearsal of Address Canvassing Revealed Persistent Deficiencies in Approach to Updating the Master Address File*, OSE-18599. Washington, D.C.: Department of Commerce OIG.

facilitate developing a methodology to geocode these and similar types of records that will continue to accumulate in the MTdb.

Census cannot rely on end-of-decade operations to geocode addresses for the 2020 census. As the 10 million ungeocoded records identified in 2008 steadily decreased, the updating activities continued to add more ungeocoded records to the MAF. Consequently, as of June 2011, there were 3.5 million ungeocoded records. If steps are not taken to continuously improve geocoding throughout the coming decade, we expect the number of ungeocoded records to rise as it did during the last census. For the 2010 census, funding constraints canceled geocoding operations mid-decade, resulting in 10 million ungeocoded records 2 years before the decennial census.

For the 2020 census, the bureau must maintain the MTdb throughout the decade by continuously geocoding—or address lists may become inaccurate.

For the 2020 census, the bureau must maintain the MTdb throughout the decade by continuously geocoding—or address lists may become inaccurate, and the bureau will again have to rely on expensive end-of-decade operations, abandoning plans to reengineer the census and realize costs savings. Targeted address canvassing can only occur if the MTdb is updated and accurate. Census should conduct the necessary research, develop a proven methodology, and allocate the necessary funds to continuously improve address quality and reduce the number of ungeocoded addresses throughout the decade.

III. Census's Process of Updating Addresses by Using the Most Recent Address May Result in a Lower Quality Address List

In our review of the procedures for updating the MTdb, we found that when different operations submit changes to the address list, Census may replace a valid address in the MTdb with one that is invalid or of lower quality. This occurs because Census accepts the latest update without determining the quality of the change. Consequently, recently collected flawed address information may replace an existing valid address.

Our review of a 2008 Census test found that changes made by a later operation were sometimes incorrect.¹² Specifically, we reviewed the actions (e.g., verify, delete, add, duplicate, uninhabitable, or nonresidential) taken by several operations for nearly 19,000 addresses from the North Carolina test site and found examples where the final action was incorrect. To analyze the impact of this during the 2010 census, we again compared the actions made to an address during the address-canvassing operation with an ACS update file. The files contained

¹²Commerce OIG, *Census 2010: Dress Rehearsal of Address Canvassing Revealed Persistent Deficiencies in Approach to Updating the Master Address File*.

the results of the two field operations conducted at the same approximate time, during the months of March through July 2009.¹³

We were unable to compare identical addresses by matching the address identification number—the ACS update file contained only the address identification number (not the complete address) with a designation of the change made (e.g., designating a residential address as nonresidential). Although our address-canvassing file was limited to 599 counties, we identified 2,457 address records contained in both files. For those addresses, we compared the actions taken by the address-canvassing staff with the actions taken by the ACS staff, to determine whether two field operations conducted at the same time could have differing results. Our analysis found agreement on the designation of the address only 56 percent of the time (e.g., both designated the address as nonresidential). We were unable to assess the level of agreement for 30 percent of the addresses because of variations in the codes used for the two operations. For the remaining 14 percent, our analysis found clear disagreement—meaning that address canvassers verified the address as residential, while ACS staff deleted or changed the address status to some other nonresidential classification (see table 2).

Table 2. Comparison of Two Census Designations for the Same Addresses

Designation by Address-Canvassing Staff	Designation by American Community Survey Staff	Comparison	Percentage (Number of Addresses)
Valid residential address	Delete, uninhabitable, nonresidential, group quarter	Disagree	14% (346)
Delete, uninhabitable, nonresidential	Delete, uninhabitable, nonresidential	Agree	56% (1,380)
Other	Other	Unable to assess ^a	30% (731)
		Total	100% (2,457)

Source: OIG

^aDifferences between the two operations made it difficult to assess the level of disagreement for some status codes. For example, address canvassing identified duplicates as a result of visiting every housing unit; ACS visited housing units based on a targeted list of nonresponding housing units and did not identify duplicate addresses. Also note that ACS updates were not entered into the MTdb because a moratorium on updates was enacted in preparation for the 2010 census.

¹³ Address canvassing occurred between March 30, 2009, and July 10, 2009; ACS (time of interview operation) took place from March 1, 2009, through May 30, 2009.

We reviewed 80 of the 346 addresses that showed clear disagreement. We do not have enough information to conclude which operation was more accurate, or why; we found that 50 addresses were included in the 2010 census (i.e., were valid housing units) while 30 were not.

Although these findings cannot be projected to the entire MTdb, they clearly indicate that field operations produce conflicting information, some of which may be incorrect. Moving forward, the bureau needs to confirm that any update made to an address is of better quality than the information currently existing in the MTdb.

Census is considering implementation of some quality indicators for addresses that might eliminate automatically accepting the most recent updates. Specifically, the following questions would be answered before a change to an existing address was made:

- Is the existing MTdb address in a valid Postal Service format (mailable)?
- Is the Postal Service delivering mail to the address (deliverable)?
- Are field staff able to navigate to, and identify/distinguish, the address (locatable)?
- Is the address linked to a location on the TIGER map (geocodable)?

We support this effort. However, we suggest that the bureau consider an additional indicator: verifiability. More specifically:

- Is the address a valid address in recent Census surveys or other government administrative records (verifiable)?

Both the bureau and OIG have identified administrative records¹⁴ as a valuable tool for future Census operations, and we believe address information contained in these data files could indicate the quality of an MTdb record. The Census Bureau has been using one form of administrative record from the U. S. Postal Service since 1995 to provide new addresses for the MAF. Data collected by Census and other agencies provide a complementary universe of information: addresses where individuals indicate they live and/or want their mail directed. Administrative records could be used to verify MAF records, leveraging the general public's notification (validation) of their residence to the government.

We suspect shifting from a last-in model will likely encounter some resistance for fear that an address will be missed. Consequently, the long-standing practice of accepting all field changes as fact may be difficult to overcome. However, to prevent inaccurate information from replacing accurate information, quality standards must be implemented.

¹⁴ Data collected for the administration of programs and provisions of services by federal, state, and local governments and commercial entities.

Recommendations

We recommend that the Director of the Census Bureau take the following actions:

1. Develop an MTdb measure for determining address list quality at a low level of geography that (a) provides a fair and equal opportunity for targeting selection, (b) drives selection and planning decisions, and (c) is well-documented and verifiable.
2. Work with the Department to determine the feasibility of improving methods of sharing MTdb information throughout the decade with governmental entities (partners) to create a uniform, national address list.
3. Investigate and remedy the exclusion of 500,000 ungeocoded address records, which had been designated as valid U.S. Postal Service delivery addresses, from the 2010 census.
4. Conduct the necessary research, develop a proven methodology, and allocate the necessary funds to continuously reduce the number of ungeocoded records throughout the decade.
5. Develop and implement quality indicator tools, including use of administrative records, to ensure that updates to the MAF are accurate.

Summary of Agency and OIG Comments

In its April 30, 2012, response to our draft report, the Census Bureau concurred with four of our recommendations and has taken some preliminary actions to address them. Although the bureau generally agreed with recommendation 2, it stated that this recommendation warranted further discussion within the Census Bureau and with the Department of Commerce rather than describing preliminary actions. In addition, the bureau stated that recommendation 3 inaccurately reported that 500,000 housing units were missed during the decennial census. We did not say that the housing units were missed but instead that 500,000 valid Postal Service address records were not included in the 2010 census. We recognize that the addresses may have duplicated an existing MTdb record that was counted, may have been demolished or destroyed during the decennial census time frame, may have been excluded due to Postal Service error, or may have been missed—hence our recommendation to investigate why the records were not geocoded during the 100 percent address-canvassing operation or included in the final decennial housing count. The bureau's response is included as appendix B.

In a separate document, the bureau provided a number of technical comments that we addressed in the report where appropriate. It also discussed the accomplishments of the 2010 MAF/TIGER Enhancement Program's (MTEP's) five objectives. MTEP had many successes and improved the overall quality of the MTdb, but it did not develop a measure to assess MTdb quality at low levels of geography.

Appendix A: Objectives, Scope, and Methodology

The objectives of this evaluation were to (1) review the bureau's progress toward achieving the objectives of its \$496 million 2010 geographic initiative, (2) evaluate the impact of various address updating operations on the MTdb and identify trends that introduced error, and (3) review Census's procedures for updating the MTdb files.

We met with and interviewed various Census headquarters personnel with broad and varied MTdb responsibilities, including those involved in its design, development, testing, and operation. We visited the Philadelphia regional office staff and conducted telephone interviews with geographers at five other regional offices. We also observed the batch update of the U.S. Postal Service's Delivery Sequence File and the interactive update of information obtained from Census's Boundary and Annexation Survey operation.

We obtained and reviewed the policies and procedures for updating the MTdb system, the controls used to protect data accuracy and integrity, and related security documentation. We did not test the operational effectiveness of security controls and policies. Rather, we reviewed the documentation for reasonableness and adequacy.

To evaluate the impact of various address-updating operations on the MTdb and to identify trends that introduced error, we requested and analyzed nine data files, including files containing MTdb updates from various 2010 census operations and Postal Service update files. Data files from two 2010 decennial ad hoc activities¹⁵ did not meet data reliability standards and, therefore, were not used.

Early in our evaluation, Census informed us that many of the 2010 geographic initiative (MAF/TIGER Enhancement Program) project managers were no longer with the bureau and that producing extensive documentation about the initiatives would be difficult and time-consuming. Given this constraint, we reviewed the information that was provided to us and focused primarily on the areas of the initiative that were not fully implemented.

We performed the evaluation at Census Bureau headquarters in Suitland, Maryland, and the Philadelphia regional office. We conducted this review from January 2011 to October 2011, under the authorities of the Inspector General of 1978, as amended; Department Organization Order 10-13, August 31, 2006, as amended; and in accordance with the *Quality Standards for Inspection and Evaluation* (January 2011) issued by the Council of the Inspectors General on Integrity and Efficiency.

¹⁵ One file came from the Address-Canvassing Delete Check operation that was implemented during the nonresponse follow-up operation. This operation determined whether address canvassers deleted housing units located in the wrong Census block but did not subsequently add the units back into the correct Census block. The second data file was created from documents collected during the address-canvassing operation, when technology barred adding housing units or correcting deleted housing units during quality control activities.


Appendix B: Agency Response




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

APR 30 2012

MEMORANDUM FOR Ann Eilers
Principal Assistant Inspector General for Audit and Evaluation

Through: Mark E. Doms 
Chief Economist

From: Robert M. Groves 
Director

Subject: Draft Report: *High-Quality Maps and Accurate Addresses Are Needed to Achieve Census 2020 Cost-Saving Goals*

The attached comments are in response to your draft audit report, *High-Quality Maps and Accurate Addresses Are Needed to Achieve Census 2020 Cost-Saving Goals*. The U.S. Census Bureau appreciates the comments and recommendations developed by the Office of the Principal Assistant Inspector General in producing this audit report.

Attachment

cc: US/EA

U S C E N S U S B U R E A U
Helping You Make Informed Decisions

www.census.gov

OIG Draft Report
High-Quality Maps and Accurate Addresses Are Needed
to Achieve Census 2020 Cost-Saving Goals

U.S. Census Bureau Response

The Census Bureau has reviewed the Department of Commerce, Office of Inspector General (OIG), Draft Report entitled “*U.S. Census Bureau High Quality Maps and Accurate Addresses Are Needed to Achieve Census 2020 Cost-Saving Goals*,” dated March 23, 2012. We acknowledge the goals of the draft report, agree with many of the points raised in the report, and agree that the report contributes to future directions in improving addresses that support cost avoidance goals of the 2020 Census.

There is, however, some information that requires further clarification, and there are other points that require correction. Our comments follow:

Comments Specific to the Recommendations:

Recommendation 1: Develop an MTdb measure for determining address list quality at a low level of geography that (a) provides a fair and equal opportunity for targeting selection, (b) drives selection and planning decisions, and (c) is well-documented and verifiable.

Census Bureau Response: The Census Bureau agrees with this recommendation. As part of the Geographic Support System Initiative, the Geography Division formed ten working groups to support research and development activities. The Quality, Assessments, and Evaluations Working Group recommended the development of Quality Indicators to measure the quality of the address list. An Implementation Project Team is currently developing the criteria and requirements associated with the Quality Indicators. In addition, after a thorough analysis, the Geography Division decided to track the Quality Indicators at the census tract level, a low level of geography that can provide sufficient detail to support decision-making.

Recommendation 2: Work with the Department to determine the feasibility of improving methods of sharing MTdb information throughout the decade with governmental entities (or partners) to create a uniform, national address list.

Census Bureau Response: The Census Bureau appreciates this recommendation and believes it warrants further discussion within the Census Bureau and with the Department of Commerce.

Recommendation 3: Investigate and remedy the exclusion of 500,000 ungeocoded address records, which had been designated as valid U.S. Postal Service delivery addresses, from the final housing count for the 2010 Census.

Census Bureau Response: It is inaccurate to state that these 500,000 housing units were missed as part of the decennial census. There are many reasons an address may be ungeocoded in the MTdb, including missing features, street names, and/or address ranges. The housing units these addresses represent may be present within the MTdb as non-city style addresses, such as location descriptions, where a city-style address could not be obtained.

The Census Bureau agrees that we need to analyze ungeocoded addresses within the MTdb. We are in the process of implementing the MAF/TIGER Address Geocoding (MTAG) program to continuously review and eliminate ungeocoded addresses throughout the decade.

Recommendation 4: Conduct the necessary research, develop a proven methodology, and allocate the necessary funds to continuously reduce the number of ungeocoded records throughout the decade.

Census Bureau Response: The Census Bureau agrees with this recommendation. As mentioned above, we are in the process of implementing the MTAG program to continuously review and eliminate ungeocoded addresses throughout the decade.

Recommendation 5: Develop and implement quality indicator tools, including use of administrative records, to ensure that updates to the MAF are accurate.

Census Bureau Response: The Census Bureau agrees with this recommendation. As mentioned above, an Implementation Project Team is currently developing the criteria and requirements associated with Quality Indicators. We will consider the recommendation to include one additional indicator (verifiable) to determine if an address is valid in recent Census surveys or other government administrative records.

I200100-119