Testimony of

Carol Rice

Assistant Inspector General
for Economic and Statistical Program Assessment

U.S. Department of Commerce
Office of Inspector General

before a hearing of the

Committee on Oversight and Government Reform
U.S. House of Representatives

The Census Bureau's Modernization Efforts and Overall Preparedness for the 2020 Census

June 9, 2016

Chairman Chaffetz, Ranking Member Cummings, and Members of the Subcommittee:

I appreciate the opportunity to testify today about the Census Bureau's modernization efforts for the 2020 Census, as well as the bureau's overall preparedness for carrying out key aspects of the 2020 Census. This testimony is informed by our oversight of last decade's decennial—and, primarily, the bureau's planning to date of the 2020 Census.

Plans to automate 2010 Census field data collection had to be greatly curtailed late in the decade. Problems developing and operating the handheld computers and related automation compelled the bureau to abandon its plan to use the devices during nonresponse followup (NRFU) and forced it to make late-stage preparations instead for a pen-and-paper operation. This change led to major cost escalation, disruption of workflow, and high operational risk. For the 2020 Census, the bureau is committed to conducting a decennial census at a lower cost per household (adjusted for inflation) than the 2010 Census.

Our testimony today will address the bureau's 2020 Census efforts toward

- accurately accounting for and estimating 2020 Census costs;
- containing costs while maintaining accuracy through an innovative decennial design; and
- ensuring preparedness through project planning and risk management.

1. Accurately Accounting for and Estimating 2020 Census Costs

Throughout the decade leading up to the 2010 Census, the Census Bureau remained uncertain of what the 2010 Census' total cost would ultimately be. With a life-cycle cost estimate of more than $11 billion established in 2003 that eventually topped $14 billion in 2008, the 2010 Census ultimately totaled in excess of $12 billion. The final cost was nearly twice the cost of the 2000 Census (in nominal dollars)—due in part to a late-stage design change and higher-than-expected contractor costs. In recent history, the cost of the decennial census has roughly doubled during each cycle.

Over the last 3 years, OIG has issued a number of findings and recommendations to the bureau related to 2020 Census cost accounting and estimates, with which the Census Bureau has concurred. Some recurring themes in our reports include recommendations to improve controls over cost monitoring and budget development.

Accounting for actual costs. Our May 2014 report, *The Census Bureau Lacks Accurate and Informative Cost Data to Guide 2020 Census Research Through a Constrained Budget Environment,* identified control weaknesses in the Decennial Program's method for recording salary costs. Neither specific project costs nor the cost of the entire Decennial Program's research effort to date could be determined, because project costs were recorded in the accounting system simply to match previously set budget allocations. Additionally, recorded salary costs did not necessarily account for what the employee actually worked on, so some projects could have been charged to incorrect activities and accounts. Inadequate accounting of employees' actual work and level of effort required to accomplish project goals and inaccurate project costs hinder the Decennial Program's ability to assess the return on investment of research efforts. It

also makes it difficult to make informed decisions about how to implement budget reductions. In addition, we found that the bureau's Decennial Program was unable to provide us with documentation supporting fiscal years (FYs) 2013 and 2014 Congressional budget justifications by tying specific requests in the President's budget to specific project activities.

The bureau's FY 2013–2015 budget justifications for 2020 Census research and testing (R&T) program assumed $5.2 billion in savings if design innovations could be implemented. Likewise, a regional office realignment was justified by an annual savings estimate of $14–$17 million per year, beginning in FY 2014. However, recently we identified issues with the bureau's cost estimates, and the bureau has not been able to demonstrate that actual cost savings can or will be achieved.

To effectively manage a program of the size, complexity, and cost of the 2020 Census—and assess the return on investment of R&T—managers need to develop detailed and supportable cost estimates to use as benchmarks for success. The estimates should then be compared to actual costs, to assess the return on investment of R&T. The bureau must improve its cost estimation and accounting practices in order to provide stakeholders assurance that budget requests are justified and will yield expected results.

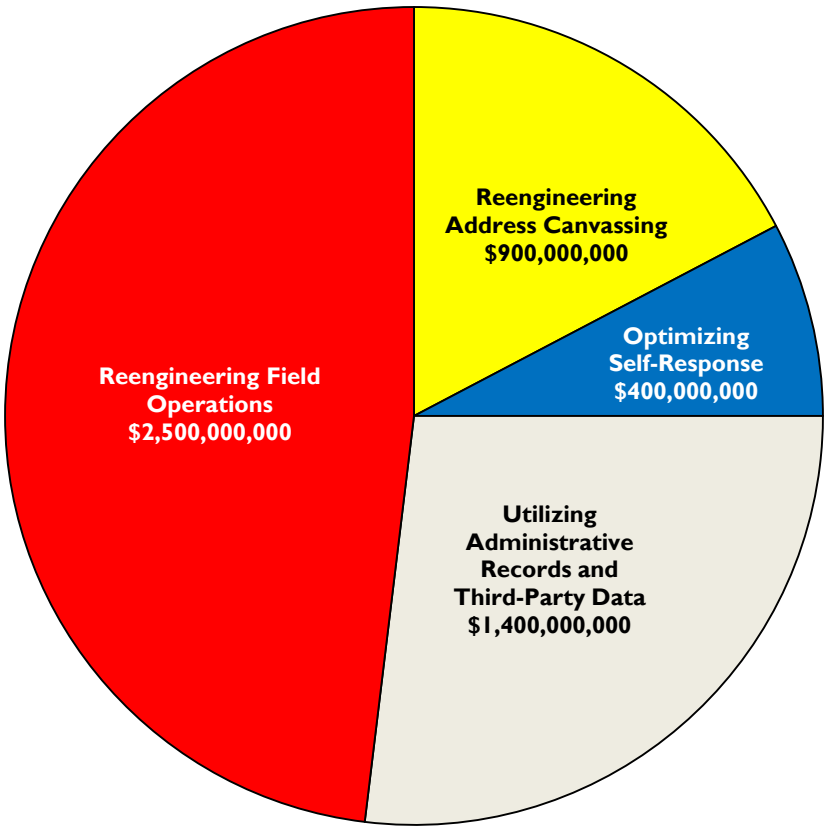
Supporting cost estimates. In our September 2015 report, *2020 Census—The 2014 Census Test Misses an Opportunity to Validate Cost Estimates and Establish Benchmarks for Progress,* we noted that the 2014 Census Test was a missed opportunity to validate cost estimates. We found that (1) the bureau's cost estimate lacked adequate documentation, (2) project teams did not follow project plan management and change control protocol, (3) the cost estimate did not account for some design features that were included or dismissed as viable options for the 2020 Census design, (4) the 2014 test did not provide cost data that can be used to validate cost savings estimates or compare the cost of various design strategies under consideration, and (5) 2014 test projects did not develop measurable success criteria, with which to validate potential cost savings or establish benchmarks for a cost-benefit analysis of test results.

In February and May of 2016, we issued two reports resulting from one audit on the Master Address File (MAF)/Topologically Integrated Geographic Encoding and Referencing system (TIGER) database (MTdb) and Local Update of Census Addresses (LUCA) for the 2020 Census. The February 2016 audit report, *The U.S. Census Bureau's Efforts to Ensure an Accurate Address List Raise Concerns over Design and Lack of Cost-Benefit Analysis*, aimed to (a) determine how efforts, such as the 2015 Address Validation Test, support the accuracy of the MAF and (b) evaluate the preparation of the LUCA program. Again, we found that the bureau did not collect cost data or conduct a cost-benefit analysis for two canvassing tests. And the May 2016 audit report, *The U.S. Census Bureau Geography Division Lacks Complete Information for Project Costs and Has Not Fully Monitored Geographic Support System Initiative (GSS-I) Goals*, found that the bureau did not identify costs for projects associated with continuously updating the MTdb.

2. Containing Costs While Maintaining Accuracy Through an Innovative Decennial Design

For the 2020 Census, the Census Bureau is expected to conduct a decennial census at a lower cost per household (adjusted for inflation) than it did during the 2010 Census—and maintain or improve the quality of the data it collects during the 2020 Census. To accomplish these goals, the bureau must overcome many challenges such as a constrained fiscal environment, rapidly changing technology, declining public participation, growing distrust in government, increased population diversity, complex living arrangements, and increased population mobility. To overcome these challenges and accomplish its 2020 Census goals, the bureau is focusing on four key innovation areas: (1) reengineering address canvassing; (2) optimizing self-response; (3) utilizing administrative records and third-party data; and (4) reengineering field operations. (Please see figure 1, below, for projected cost savings of each area.)

Figure 1. 2020 Census Innovation Areas and Projected Cost Savings



Source: U.S. Census Bureau, *2020 Census Operational Plan*, December 2015

The 2020 Census planning effort involves 34 operations. Of those, the address canvassing and NRFU operations are likely the most expensive; the bureau expects innovations in each of these four areas to reduce the cost of 2020 Census fieldwork by $5.2 billion (compared to the 2010 Census).

Reengineering address canvassing. By reengineering address canvassing, the Census Bureau expects to reduce the number of addresses needing field verification by 75 percent and save $900 million. To accomplish this, the bureau is focused on updating MTdb throughout the decade and conducting in-office and targeted in-field address canvassing operations.

Updating Census addresses and maps. The Census Bureau must enumerate the population in a decennial census that associates all addresses with a geographic location. The bureau therefore attempts to locate all MAF addresses geospatially in its TIGER system. As the backbone of the bureau's survey operations, the MTdb must be up-to-date and accurate. Because there is no single source for updating data in the MTdb, the bureau coordinates with providers of multiple data sources, such as tribal, state, and local governments; conducts its own operations to verify and update addresses and maps; and receives updates twice a year with the Delivery Sequence File (DSF) from the U.S. Postal Service.

After the 2010 Census, the bureau determined a need for a larger-scale effort for continuous MTdb validation and updates. This prompted the initiation of GSS-I, a continuous plan on a 10-year life cycle to provide the most current, accurate, and complete address, feature, and boundary data. To date, we have reviewed the LUCA and GSS-I programs. We found that the bureau eliminated two of the three 2010 LUCA participation options for the 2020 LUCA—and may not provide adequate alternatives for governments who find the remaining option too burdensome. As a result of the bureau's decision, the only participation option for the 2020 LUCA program participants is to identify and submit discrepancies between their residential address lists and the bureau's master address list.

The bureau stated that GSS-I could help government entities affected by the change. However, in our audit of the GSS-I program, we found the bureau did not monitor over one-third of the original goals it identified in its operational plan. As a result, the bureau does not know whether it has fulfilled those goals. Furthermore, with less than 4 years before the next decennial census, there may not be enough time to include all 2010 LUCA participants in the GSS-I program.

Reducing in-field address canvassing. In 2009, Census Bureau canvassers traversed almost every block in the nation to validate the address list. For the 2020 Census, the bureau is replacing 100 percent in-field address canvassing with a 100 percent in-office review and reduced in-field canvassing. To support this, the bureau has begun an operation that uses imagery to identify geographic changes that require further office review to update the MTdb and pinpoint areas for in-field verification. The in-office imagery review has been in place for over 6 months. We are currently assessing this operation to see whether it is being implemented as planned.

Optimizing self-response. The goal of optimizing self-response is to communicate the importance of the 2020 Census to the U.S. population and generate the largest possible self-response, reducing the NRFU workload. By offering an Internet response option,

the Census Bureau hopes to enable people to respond from any location at any time and via multiple types of electronic devices.

The bureau also hopes to allow people to respond without a unique identification code (i.e., non-ID processing), easing response to the 2020 Census. To accomplish this, however, the bureau must figure out how to collect address information and match it, in real time, to the MAF/TIGER system. This requires real-time map interface and response validation—both during the response and via back-end processing. Without automated, real-time matching and processing, manual matching and geocoding is required when automated non-ID processing cannot determine an acceptable geographic location match.

Although the bureau did not implement an Internet response option for the 2010 Census, it has confirmed the feasibility via the American Community Survey, the 2012 National Content Test, the 2014 and 2015 Census Site Test, and the 2015 Optimizing Self-Response Test. The bureau has decided to provide an Internet self-response option for the 2020 Census. However, a number of challenges remain—such as how to achieve the 55 percent response rate, which the bureau requires to attain its estimated cost savings. There are also a number of additional issues regarding optimized self-response in general, and the Internet self-response option in particular, that the bureau still has to resolve before deciding on the final design of the 2020 Census.

Utilizing administrative records and third-party data. Greater use of administrative records offers the potential to enhance the decennial census in a number of important areas: from improving the MAF to finding households or individuals who may otherwise be missed to providing quality control for the enumeration process. These personal records contain information that individuals have already provided to various government agencies or commercial entities, such as their names, addresses, age, sex, race, and a wide variety of demographic, socioeconomic, and housing information.

As indicated in the Census Bureau's 2020 business plan, supplementing decennial operations with information from these records could potentially reduce enumeration costs and help the bureau avoid inaccurate enumerations by

- improving the address list;
- supplying answers to questions with invalid or blank responses;
- providing information for households that do not respond to the questionnaire, an in-person visit, or a phone interview;
- validating respondent information; and
- helping assess overall decennial accuracy (i.e., providing coverage measurement).

However, relevant statutes governing other federal agencies do not facilitate the use of administrative records by the Census Bureau because these statutes either do not compel agencies to provide their records to the bureau in response to requests or state that agencies are only required to provide certain information to the bureau.

Reengineering field operations. The Census Bureau's plans include (1) providing automated field data collection solutions to assign and monitor work, collect and submit data, and perform administrative functions such as payroll and communications; (2) streamlining (and thereby reducing) the number of staff and field offices; and (3) employing an enterprise-wide data collection and processing infrastructure. The goal of this innovation is to use new technology and automated systems to allow enumerators, supervisors, and managers to be more productive and efficient—resulting in a more effective NRFU operation and a cost reduction of $2.5 billion.

Automated field data collection. For the 2010 Census, the bureau had planned to reduce the costs of field operations by using custom mobile handheld computing devices—equipped with global positioning system capabilities—to automate the workload assignment, data collection, and information processing functions. However, the project experienced constant setbacks, including technical problems, escalating costs, and missed deadlines. In April 2008, the decision was made to abandon the plan to use the handhelds for NRFU, although they were successfully deployed for the address canvassing operation.

For the 2020 Census, the bureau—using off-the-shelf devices—has developed in-house applications that have enumerated households in the 2013–2015 site tests. We are examining the 2015 Census Test of NRFU in Maricopa County, Arizona, to assess (a) whether the bureau's reengineered and automated operational control system for managing fieldwork functioned as expected and (b) the bureau's progress for determining whether enumerators are able to use personally owned mobile devices to collect household data, as well as the status of the bureau's efforts to overcome policy and legal issues associated with the use of those devices.[1]

Streamlined 2020 Census field structure. By providing enumerators with the capability to perform all administrative and data collection tasks remotely—directly from a handheld device—and enabling supervisors to work remotely and communicate with their staff via the devices, the bureau expects these enhanced capabilities to significantly reduce the number of local field offices required to support 2020 Census fieldwork. The bureau also expects this increased automation to make it easier for supervisors to monitor and manage their enumerators. Therefore, the bureau assumes that the ratio of enumerators to supervisor can be increased, reducing the number of supervisors required.

We are also reviewing the 2016 Census Test of NRFU in Harris County, Texas, and Los Angeles County, California, to assess whether—through (a) increased use of technology, (b) streamlined office and staffing structure, and (c) increased

[1] In January 2016, the Census Bureau decided to eliminate "Bring Your Own Device" as an option for providing enumerators with devices or smartphones. Instead, the bureau decided to implement the Device as a Service (DaaS) strategy for providing enumerators with equipment during the 2020 Census. Under the DaaS option, a single vendor—at a single cost—will supply the necessary equipment, handle all logistics, configure the devices, manage inventory, and provide technical support.

management and staff productivity—the bureau will be able to manage more enumerators with fewer supervisors during the 2020 Census. During the 2016 test, the bureau is assessing two different staffing ratios—20 enumerators to 1 supervisor in Los Angeles County, and 30 enumerators to 1 supervisor in Harris County. This is a significant increase from the number of enumerators (approximately 8) each supervisor managed during the 2010 Census. Our fieldwork in this area is ongoing, and no findings have been developed. At the end of the test, the bureau intends to assess the cost and quality tradeoffs associated with each new, greater staffing ratio.

Census Enterprise Data Collection and Processing (CEDCaP) program. The bureau's goal is to have mature, proven systems in place well in advance of the 2020 Census to avoid building one-time use applications. To accomplish this, the bureau intends to deliver an integrated and standardized network of systems, using an enterprise-wide approach to collect and process data through shared services.

However, in the past, the bureau has struggled with program management and meeting scheduled benchmarks for its information technology (IT) development programs. With less than 4 years remaining until Census Day (April 1, 2020), the bureau is running out of time to successfully deploy an enterprise solution. Yet the bureau must develop, test, and implement a cost-effective, secure IT infrastructure to collect and process data in time to support the 2020 Census workload volume.

As part of our FY 2016 audit plan, we recently initiated an audit of the planning, design, incremental development, and implementation of CEDCaP. Our objectives are to (1) determine whether applicable enterprise architectural and system development standards, methodologies, and best practices are being effectively employed on the CEDCaP program, and (2) assess the adequacy of plans for scalability, reliability, and security throughout the development lifecycle of CEDCaP cloud-based applications.

3. Ensuring Preparedness Through Project Planning and Risk Management

The Decennial Program began its 2020 Census R&T phase with a plan to conduct 24 small, medium, and large field tests to mature innovations in an iterative manner—culminating in FY 2014 with a large integration field test. However, during late 2012 and early 2013, the Decennial Program R&T activity schedule was revised four times and activities were delayed with each revision. Under this *initial* revised schedule, the R&T field tests were set to conclude in FY 2016. In June 2013, citing budget constraints, the bureau announced its plan to cut the number of field tests leading up to the large integration test to 11. The 2014 large integration test has been recast as the Census End-to-End Test—now scheduled for 2018.

The bureau only has a limited number of opportunities left to develop, test, and refine 2020 Census design alternatives, as well as resolve remaining issues. Along with the 2016 Census Test, and the 2016 Address Canvassing Test, there remains the 2017 Census Test, the 2018 Census End-to-End Test, and Post End-to-End Testing in 2019.

The 2020 Census design includes 34 operations at various planning stages. In addition to determining the final content topics and questionnaire wording for the 2020 Census, the bureau has identified 184 design issues, which still have to be resolved before the final 2020 Census design decisions can be made (see appendix A). As of June 2016, these unresolved design issues include 92 related to response data collected by enumeration operations and, of those, 20 specifically related to NRFU.

The Decennial Program—which must finalize its design decisions well before Census Day 2020—faces the challenge of completing its mission under the pressure of tight deadlines. In addition to this challenge, many R&T projects have lacked measurable success criteria with which to validate potential cost savings or establish benchmarks for analyzing test results. Many of the projects included in past tests have lacked certain requirements promulgated by GAO, as well as the bureau itself, including

- research results that address specific research questions in a timely manner;
- recommendations resulting from projects that are supported by appropriate evidence;
- outputs—such as data, products, or objects—during the project's life cycle;
- performance measures to evaluate progress towards achieving the outputs and to evaluate project and program performance against pre-established targets; and
- performance metrics that (1) align with the goals and mission of the R&T effort and that are clearly communicated, (2) are clearly stated with a unique name and definition and include a specific methodology used for calculation, (3) have quantifiable goals that can be achieved during a defined time period, (4) are free from significant bias and produce the same result under similar conditions, (5) cover activities that support program goals and objectives, and (6) provide new and unique information.

These projects lacked success criteria that established precise, predetermined, and quantifiable benchmarks against which to (a) measure actual test results and (b) determine whether projects or tests can achieve specific goals or objectives. The lack of appropriate success criteria for R&T projects—as well as the lack of corresponding cost estimates—restricts the bureau's ability to base decisions on actual results and inhibits the demonstration that expenditures produced quantifiable results that will help the bureau conduct a cost-effective and quality 2020 Census.

I will be pleased to take your questions.

Appendix A.
2020 Census Design Decisions Yet to Be Resolved

Area Operation	Oct 2015	Nov 2015	Dec 2015	Jan 2016	Feb 2016	Mar 2016	Apr 2016	May 2016	Jun 2016	Jul 2016	Sep 2016	Oct 2016	Dec 2016	Jan 2017	Mar 2017	Apr 2017	Jun 2017	Aug 2017	Sep 2017	Oct 2017	Nov 2017	Dec 2017	Jan 2018	Apr 2018	Sep 2018	Dec 2018	Mar 2019	Jun 2019	Jun 2021	Undefined	Grand Total
Census/Survey Engineering						1			1		1	1				1			2					1							8
Content and Forms Design												1				1			1					1							4
Language Services																			1											1
Security, Privacy, and Confidentiality						1																								1
Systems Engineering and Integration									1		1																			2
Frame			1	1						1		2		5	4		2	1		2											19
Address Canvassing														5	3															8
Geographic Programs				1											1		1	1		2										6
Local Update of Census Addresses			1							1		2					1													5
Infrastructure			2		2	1			2		2			5						1											15
Decennial Logistics Management			1			1																								2
Decennial Service Center					2									3																5
Field Infrastructure			1											2																3
IT Infrastructure									2		2									1											5
Other Censuses																			3			1								4
Island Areas Censuses																			3			1								4
Program Management																													8	8
Program Management																													8	8
Publish Data		1								1		1			1	1	1		10			1		2		1	1	1		22
Archiving																			3			1								4
Count Question Resolution																								2				1		3
Count Review												1							7											8
Data Products and Dissemination		1								1					1															3
Redistricting Data Program																1	1									1	1			4
Response Data	1	2	7	1		2	1	1	7		12	6			1	3			21	11	1	7	4		4					92
Census Questionnaire Assistance		2					1		2		1					1					1		2							10
Enumeration at Transitory Locations																			6											6
Federally Affiliated Americans Count Overseas																							2							2
Forms Printing and Distribution	1											2								1										4
Group Quarters									1											1		6								8
Integrated Partnership and Communications															1	2														3
Internet Self-Response				1								1							1	4										7
Non-ID Processing											2								2					2						6
Nonresponse Followup											9								9					2						20
Paper Data Capture												3								1										4
Response Processing			3					1											1											5
Update Enumerate			4			2			4										2	4		1								17
Test and Evaluation			1			1					8		2						3							1				16
Coverage Measurement Design and Estimation			1			1					5								1											8
Coverage Measurement Field Operations											2								1											3
Coverage Measurement Matching											1								1											2
Evaluations and Experiments													2													1				3
Grand Total	1	3	11	2	2	5	1	1	10	2	23	10	2	10	6	5	3	1	39	14	1	9	4	1	6	1	1	1	1	8	184

Source: U.S. Census Bureau, *2020 Census Operational Plan*, December 2015