Written Testimony
of
John H. Thompson
Director of the U.S. Census Bureau

Before the U.S. House of Representatives
Subcommittee on Government Operations
and the
Subcommittee on Information Technology
of the Committee on Government Oversight and Reform

November 3, 2015

Chairman Meadows, Chairman Hurd and Ranking Members Connolly and Kelly, thank you for the opportunity to update the House Subcommittees on Government Operations and Information Technology on the U.S. Census Bureau's progress in planning for the 2020 Census. I am also pleased to be testifying with Mr. Steve Cooper, the Chief Information Officer of the Department of Commerce with whom we work closely.

I am honored to serve as the Director of the Census Bureau, which has the Constitutional responsibility of administering an accurate census that fairly represents every person living in America, and is fundamental to our democracy. We have reached a major milestone in our preparations for the 2020 Census. We have released an operational plan that allows us to change fundamentally the way in which the Census Bureau has conducted the Decennial Census for over 40 years. The 1970 Census was a breakthrough for its time. We built an address list and mailed questionnaires to each housing unit on the list. We asked respondents to complete and return the questionnaires through the mail. We developed automated processes to capture the information on the returns. However, the task of collecting information from those households that did not self-respond required recruiting and managing a army of enumerators using paper and pencil. For each census since 1970, this paper-based process has been the standard, and it has been increasingly challenged by the growing diversity and complexity of our nation. We do not believe that a paper and pencil approach to the Census is sustainable for the 2020 or future censuses. It is no longer affordable and more importantly, is not adequate to meet the challenges of enumerating an increasingly diverse society. The operational plan that we prepared incorporates a number of innovations that will allow us to achieve a modern census for the 21st century. We are proposing to use mobile technology, administrative records, innovations from the geospatial industry, and self-response via the Internet to build a census that will cost far less than repeating the outdated processes used in 2010 to accurately count every person living in America.

I will briefly discuss the components of our 2020 Census Operational Plan to provide an overview of the key operations and innovations it includes. We are also pleased that this plan has been released three years earlier than the corresponding plan for the 2010 Census. It includes a number of important decisions that are based on the important research we have conducted and clearly documents the process by which we will conduct additional development and testing to have all key decisions in place to support a complete end-to-end test of our methodologies in 2018.

The Census Bureau has concentrated on four key innovation areas: reengineering address canvassing; optimizing self-response; utilizing administrative records and third-party data; and reengineering field operations. Each of these combined innovation areas has the potential to dramatically improve census operations and to save taxpayers money. At this time, the 2020 Census is estimated to cost $12.5 billion compared with a cost of $17.8 billion that it would take to repeat the paper and pencil design of the 2010 Census. This represents approximately $5.2 billion in cost avoidance.

I will briefly describe each of these innovation areas.

First, we must make sure we have a complete address list, and to accomplish this, we will canvass the entire nation as we have in each previous census. However, for 2020 we have reengineered the process to incorporate a one hundred percent in-office canvassing that will be supplemented with a twenty-five percent field operation. We have already started this effort and we expect to realize $900 million in cost savings from this strategy.

Second, to generate the largest possible self-response, our strategy incorporates not only the Internet as the principal response option, but also a nationwide integrated communication and partnership strategy and tailored contact strategies. We will mail an invitation to all housing units, encouraging Internet response. For 20 percent of the addresses in areas with low Internet connectivity and older populations, we will be including a questionnaire in the first mail package. We will also incorporate a Census Questionnaire Assistance program with both telephone and web-chat options. We will make it easy for people to respond anytime and anywhere, which means we will accept responses over the Internet, even when respondents cannot locate and provide their unique census identification code. We believe that by opening up the process and making it easy, we will encourage participation with respondents who have typically not responded. These efforts offer the potential of approximately $400 million in savings.

Third, using administrative records and third-party data will also increase our efficiency by reducing the workload that we face in collecting information from those housing units that do not provide a self-response – an operation we refer to as nonresponse follow-up. Our current estimate for the non-response follow-up workload for the 2020 Census is 56 million housing units. To reduce the workload, we are proposing to use administrative records for enumeration in two ways.

We are planning to use administrative records to remove the vacant addresses before sending census takers into the field, meaning enumerators will not knock on doors where no one is living, which represents a significant number of addresses, approximately 11 percent of the workload or six million housing units from the total workload of 56 million. We are also planning to use administrative records to enumerate occupied housing units after attempting to obtain a response through an in person visit. We will knock on every door at least once, which we anticipate will net at least 11 million additional responses, reducing the non-response workload to approximately 40 million. The next step we propose is to use administrative records from trusted federal and state sources to enumerate an additional 6 million housing units. We would then visit the remaining housing units to complete the enumeration. According to our estimates, using administrative records in this way will result in a savings of $1.4 billion.

We have identified the core set of records we will use for the 2020 Census, but we continue to assess other federal and state sources. However, one of the most important sources, the National Directory of New Hires (NDNH), is currently unavailable for most evaluation and statistical purposes government-wide, including the decennial census, because access to these data for these purposes is not currently permitted under the Social Security Act. This database supplies information on workers, including the newly hired, which the Census Bureau could use to corroborate and supplement other information. Last year, legislation permitting broadened access for Federal evaluation and statistical purposes, including the decennial census, was introduced in and partially moved by Congress. In addition, we developed draft language at the request of the Senate Committee on Homeland Security and Government Affairs. The language provided limited access for the Census Bureau to use these data for statistical purposes and ensured protection of those data under the Census Law. The President's FY 2016 Budget also highlights our need for these data.

Finally, we have reengineered our nonresponse follow-up operation to save an estimated $2.5 billion. We will now use mobile technology and smartphones to achieve significant efficiencies for the 2020 Census. I should note that we have successfully developed an innovative prototype system that incorporates commercial off-the shelf technology and software to deploy our workforce. We are now able to provide optimized work assignments to our enumerators, including daily route assignments and the best time of day to attempt contact. We are also able to provide the supervisors of our enumerators with real-time updates and alerts regarding the progress of the workers that they oversee.

Supporting these efforts is the Census Enterprise Data Collection and Processing (CEDCaP) effort, which is an enterprise approach to survey and census data collection and processing through shared services. In the past, duplicative systems were created and used for every survey and census. This duplication meant there were more than 100 systems the Census Bureau used to collect and process data, some of which, in the case of the Decennial Census, were decommissioned after use. This stove-piped approach was not the most economical or productive way to do business and the Census Bureau is working toward a sustainable, enterprise approach that simplifies and integrates data systems across the lifecycle from survey design through instrument development, survey, data collection, and data processing, editing, imputation, and estimation.

From dozens of unique systems, we will move to a small suite of shared, reusable systems with the goal of conducting a modernized 2020 Census. Based on my experience, in overseeing the 2000 Census and in the private sector, I am confident we are on course and we have a schedule for the major decisions. Over the course of the last year, we made substantial progress and developed prototype systems to eliminate the paper and pencil processes for the Decennial Census. We established milestones for moving forward and we have engaged Carnegie Mellon's Software Engineering Institute to help us make the critical decisions regarding deployment of production systems.

As a final point, I must mention our commitment to protecting the privacy and confidentiality of individuals' information. The Census Bureau takes cybersecurity seriously. Protecting privacy and confidentiality are central to the Census Bureau's mission and is a core value of our agency—we know that if the public does not trust us to protect their information, they will not provide their information to us. To protect our information systems and the information we collect, the Census Bureau has implemented a robust, comprehensive, and layered cybersecurity program. Cybersecurity is an ongoing process and challenge, and we look to outside experts to constantly evaluate our posture and our security program. Some of the key points of the program are:

- The Census Bureau utilizes the Department of Homeland Security's resources to protect Internet traffic, by looking for malicious code and suspicious activity.
- Other key safeguards include 2-factor authentication; use of encryption in transmissions of data and data at rest; use of a Data Loss Prevention System; use of firewalls, intrusion detection systems and intrusion prevention systems; a dedicated cybersecurity staff that monitors these systems and investigates unusual activity; employee fingerprint and background checks for employees, including enumerators; and all employees receive annual training on data stewardship and security.
- All Census Bureau systems are compliant with the National Institute of Standards and Technology (NIST) Federal Standards and Guidelines, and the Census Bureau security program incorporates continuous monitoring of all IT systems. We have implemented a Risk Management Framework (RMF) program, that is fully compliant with NIST standards and guidelines, and for which the initial step is to identify risks.
- Finally, we work closely with the Department of Commerce Chief Information Officer and Office of Security. We also work with our oversight bodies, the Government Accountability Office (GAO) and Office of the Inspector General, whose recommendations we take seriously.

From CEDCaP and cyber security to the 2016 Census Test and the 2018 Census End-to-End Test, each activity plays a significant role in helping the Census Bureau to design an accurate and cost-effective 2020 Census. The 2018 Census End-to-End Test represents the culmination of research and testing, as we implement the planned census operations in real-time. Even though 2018 may be two years away, we have a sense of urgency right now about what we can accomplish this year to prepare for this test. We cannot risk a major IT system failure in 2020, therefore the 2018 Census End-to-End test is critical and must have all of the major systems in place for this test. The President's Budget includes adequate funding to accomplish our objectives and to conduct a full scale 2016 Census Test, which will provide insights and guide our planning to ensure an accurate census.

The last five years have provided the groundwork for the 21st century census and now, looking forward, we must turn our attention to counting every person in America. But I must also be candid about the next five years and the risks we face. A significant risk we face to executing a 2020 that reduces cost while maintaining quality is receiving adequate funding for the entire lifecycle. Stated simply, if the funding appropriated for each fiscal year during the 2020 Census lifecycle is less than requested, then the Census Bureau must reprioritize its activities, which will affect our ability to reengineer systems and operations to conduct an accurate census. We are fully committed to a census that fairly represents every person in America. We are confident that with the President's Budget request we can build a sustainable census and count every person; however, if we have to defer activities to later years, the costs of an accurate census will escalate and the projected savings will diminish.

I am confident the Census Bureau can achieve these objectives given Congressional support and I look forward to discussing the 2020 Census Operational Plan and other aspects of our planning with you.

Thank you.