Plans and innovations for the 2020 decennial census of the United States

Lisa M. Blumerman\textsuperscript{a}, Deirdre Dalpiaz Bishop\textsuperscript{b} and James L. Dinwiddie\textsuperscript{a,*}

\textsuperscript{a}Decennial Census Programs, United States Census Bureau, Washington, DC, USA
\textsuperscript{b}Decennial Census Management Division, United States Census Bureau, Washington, DC, USA

Abstract. The purpose of the 2020 Census is to conduct a census of population and housing and disseminate the results to the President, the states, and the American people. The goal of the 2020 Census is to count everyone once, only once, and in the right place, and the challenge is to do this at a lower cost per household than the 2010 Census, while maintaining high quality results. The Census Bureau addressed this challenge first by identifying the major cost drivers of the decennial census. From 2013 through 2015, the Census Bureau conducted research and testing related to major innovations that showed promise of significant cost savings. Across four key innovation areas, the Census Bureau believes it can avoid over $5 billion in costs relative to the cost of repeating the 2010 Census design and operations in 2020. The Census Bureau used the results of the research and testing, other key information, and input from a wide variety of stakeholders to design the 2020 Census. The Census Bureau documented and published its 2020 Census Operational Plan on October 6, 2015, accompanied by revised lifecycle costs estimates and the overall 2020 Census lifecycle budget.

Keywords: Decennial Census, planning, innovations, cost savings, administrative records, data collection technology, Internet data collection, address list

1. Introduction

In the next census of the United States, the Census Bureau will be reaching out to an increasingly diverse and growing population of around 330 million people in more than 140 million housing units.

Over the past four decades, reducing the undercount has been a major goal for the decennial census. For the 2010 Census, the estimated net undercount was nearly zero – less than one half of one percent \cite{1}. The Census Bureau is committed to maintaining quality results for the 2020 Census, but there are a number of challenges it faces:

- **Constrained fiscal environment**: Budget deficits place significant pressure on funding available for the research, testing, design and development work required for successful innovation.
- **Rapidly changing use of technology**: Stakeholders expect the decennial census to use technology innovation, yet the rapid pace of change makes it challenging to plan for and adequately test the use of these technologies before they become obsolete.
- **Information explosion**: Rapid changes in information technology create stakeholder expectations for how the Census Bureau interacts with the public to obtain and disseminate data products.
- **Declining response rates**: Response rates for Census Bureau surveys, and for surveys and censuses in general, have declined. Citizens are overloaded with requests for information and have become increasingly concerned about sharing information.
- **Distrust in government**: Concerns continue to grow about information security and privacy, the confidentiality of information given to the government, and how government programs will use the information it collects. This makes it more diffi-

\textsuperscript{1}Any views expressed on statistical, methodological, technical or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

\textsuperscript{*}Corresponding author: James L. Dinwiddie, United States Census Bureau, 4600 Silver Hill Road, Washington, DC, USA. Tel.: +1 301 763 1346; E-mail: james.l.dinwiddie@census.gov.
cult to collect important demographic survey information.

- **Increasingly diverse population**: The demographic and cultural make-up of the United States continues to increase in complexity, resulting in a growing number of households and individuals who do not speak English as their native language, who have a wide variety of cultural traditions, and who may have varying levels of comfort with government involvement.

- **Informal, complex living arrangements**: Households are becoming more diverse and dynamic, making it a challenge to associate an identified person to a single location. For example, children with divorced parents may have two primary residences. Additionally, some households include multiple relationships and generations.

- **A mobile population**: The United States continues to be a highly mobile nation as about 12 percent of the population moves in a given year, based on results from the American Community Survey conducted in 2012–2013 and 2013–2014. Continued growth in the use of cellular telephone technology and an associated reduction in landline telephones tied to physical locations may also complicate enumeration.

As it becomes more challenging to locate individuals and solicit their participation through traditional methods, the Census Bureau must, decade after decade, spend more money simply to maintain the same level of accuracy as in previous censuses. On average, the total cost (in constant dollars) of conducting the decennial census has nearly doubled each decade. Initial estimates for expected total costs for the 2020 Census would be approximately $17.8 billion in 2020 constant dollars if the Census Bureau repeats the 2010 Census design and methods. With the innovations described in the 2020 Census Operational Plan [2], the Census Bureau estimates that it can conduct the 2020 Census for approximately $12.5 billion in 2020 constant dollars, thus avoiding over $5 billion in costs.

Field costs associated with Address Canvassing and Nonresponse Followup operations comprised the most expensive parts of the 2010 Census. Analyses into the operations and activities associated with these cost drivers resulted in identifying key innovation areas for 2020.

From 2012–2015, efforts focused primarily on technological and methodological innovations that provided the greatest opportunity to address the major cost drivers, enhance our understanding of the response process, and reinforce the quality of the Census. The Census Bureau developed a research and testing plan that supports the key design decisions documented in the 2020 Census Operational Plan.

During this research and testing period, the 2020 Census team focused on four key innovation areas: Reengineering Address Canvassing; Optimizing Self-Response; Utilizing Administrative Records and Third-Party Data; and Reengineering Field Operations.

All four innovation areas are aimed at maintaining the quality of the Census, while taking advantage of new data and methodologies to reduce the costs of fieldwork. A reengineered Address Canvassing operation is expected to reduce the in-field workload for address updating by 75 percent. Self-response innovations, which are aimed at generating the largest possible self-response rate, coupled with the use of administrative records and third-party data, are intended to reduce the field workload associated with Nonresponse Followup. Finally, the reengineered field operations are intended to increase the efficiency of those operations, allowing managers and fieldworkers to be more productive and effective. Each of these areas are described in more detail below.

Over the course of the next five years, from 2016 to 2020, the Census Bureau has charted a path for testing, building and integrating all components needed to conduct the census. The Census Bureau will continue to refine field operations, build and test Information Technology systems, and then integrate all parts of the plan in time to conduct an end-to-end test in 2018.

2. **Reengineering Address Canvassing for the 2020 Census**

The Census Bureau will conduct a 100 percent in-office canvass of the nation’s addresses, and will continually update the address list based on data from multiple sources, including the U.S. Postal Service, tribal, state, and local governments, satellite imagery, and third-party data providers. This office work also will determine which parts of the country still will require fieldwork because address updates cannot be obtained or verified, or because they are areas undergoing rapid change. The Census Bureau predicts in-field canvassing will be needed for approximately 25 percent of the total number of addresses (estimated at 142.9 million). These design changes have the potential to save the Census Bureau an estimated $900 million [2].

2.1. **History**

From the first U.S. census in 1790 until 1960, census
enumerators traveled from house to house while collecting housing and population information on census schedules [3]. For the first time, in 1960, the Census Bureau mailed out questionnaires to some addresses (those in urban areas) by having the U.S. Postal Service deliver forms to each address in these areas. Householders were asked to complete the questionnaire and hold it for an enumerator to collect. Enumerators compiled the address list at the same time as these visits, and also enumerated any addresses they found that had not received a form from the U.S. Postal Service.

In 1970, the Census Bureau expanded the use of mailout to about 60 percent of the population, and also employed a mailback strategy (rather than having enumerators pick up the completed forms). This required development of an address list before the census to use in labeling questionnaires and for controlling the fieldwork for visiting those households that did not respond by mail. For most urban areas, the Census Bureau built this address list by purchasing commercial address lists, followed by a U.S. Postal Service review, and then by an operation called Precanvass where enumerators walked each of these blocks to check the list for completeness. For more rural areas, enumerators canvassed each block and compiled the address list from scratch (Prelist). A similar process was used to construct the address list for the 1980 and 1990 Censuses.

In preparation for the 2000 Census, the Census Bureau started with the final 1990 Census address list rather than purchasing a commercial list [4]. “In its earliest state, the MAF (Master Address File) was an amalgamation of the 1990 Census Address Control File and the United States Postal Service’s (USPS) Delivery Sequence File (DSF). Census 2000 was the first major effort undertaken to update and, in a sense, validate the MAF” [5]. The Census Bureau used the DSF as a primary source to enhance the initial MAF, and subsequent DSF releases were used to update the address list through April of 2000. The Census Bureau also conducted Block Canvassing to update the list in more urban areas. A separate operation – Address Listing – was used to verify and, as needed build the address list from scratch, in more rural areas.

Between 2000 and 2010, the Census Bureau continued to maintain its MAF, with primary updates coming from DSF deliveries over time. With this approach, the Census Bureau was able to conduct a single operation (Address Canvassing) nationwide in 2009 to update the MAF ahead of questionnaire delivery. A major innovation for 2010 was the use of hand held computers to conduct this operation [6]. These devices contained both existing addresses and maps, along with software to manage updates, additions, and geographic locations (making use of GPS coordinates).

3. Optimizing self-response for the 2020 Census

The goal of this innovation area is to communicate the importance of the 2020 Census to the United States population, and to generate the largest possible self-response and thus reduce the workload for expensive in-person followup with those households. This will be done in several ways:

- By enabling people to respond via multiple modes (Internet, paper, or telephone);
- By allowing respondents to submit a questionnaire without a unique identification code; and,
- By providing online forms in multiple languages.

The Census Bureau plans to motivate people to respond by using technology, administrative records, and third-party data to tailor advertisements and contact strategies to different demographic groups and geographic areas. The Census Bureau also plans to utilize its partnership program for such things as providing information to trusted community leaders and spokespersons, and to host events with community, recreation, and faith-based organizations. These partnership activities, communication efforts, and contact strategies will encourage the use of the Internet as the primary response mode to a sequence of invitations and postcard mailings. These design changes have the potential to save the Census Bureau an estimated $400 million [2].

3.1. History

Since the first U.S. census in 1790, virtually all information was collected on paper, either by mail or in person. After the introduction of the World Wide Web in 1989, and the first Internet browser in 1990, the Census Bureau began to explore the use of an Internet response option [7] in the face of declining survey and census response rates [8].

The Census Bureau offered this option for the 2000 Census, but mostly due to concerns about online security, did not widely publicize the availability of Internet response. Ultimately, less than 100,000 respondents used this option in 2000 [9]. The Census Bureau studied this again for the 2010 Census, but decided it wouldn’t be cost effective to offer this option because testing had not shown an increase in overall response –
only respondents switching their mode of response [7]; see also [10].

A number of other countries began using an Internet response option over the last decade [7]. Also, in 2010, the Census Bureau began testing this option for use in the American Community Survey (ACS) [11]. The ACS is the annual survey that replaced the once-a-decade sample based collection of detailed characteristics data during the decennial census. In 2012, the Census Bureau began providing this response option for the ACS [12]. The Census Bureau also now provides this option for most of the economic and demographic surveys it conducts.

Statistics Canada used an Internet response option for their 2011 Census [13]. In their final report on plans for their 2016 Census [14], they stated: “The Internet has become the primary mode of collection for the Census Program, as almost 80% of Canadian households currently have access to the Internet. In 2011, almost 54% of Canadian households completed their 2011 Census questionnaire on the Internet, a significant increase from 18% in the 2006 Census.”

As reported by the United Nations Economic Commission for Europe (UNECE), many other countries used an Internet response option for the 2010 round of censuses [15].

4. Utilizing administrative records and third-party data for the 2020 Census

The goal of this innovation area is to use information people have already provided to improve the efficiency and effectiveness of the 2020 Census, and in particular reduce expensive in-person followup activities. Administrative record data refers to information from federal and state governments. Third-party data refers to information from commercial sources. These design changes have the potential to save the Census Bureau an estimated $1.4 billion [2]. The papers by Keller [16] and Morris et al. [17] describe some of the research the Census Bureau has done in this area.

Data from both sources can help improve the quality of the address list (frame), increase the effectiveness of advertising and contact strategies, validate respondent submissions, and reduce field workload for followup activities. As has been done in prior decades, administrative and third-party data are used to update the address frame and reflect changes that occur over time. These data also will be used to validate respondent addresses for those who respond without providing a unique Census ID. Finally, a primary use of administrative records will be to reduce the field workload for followup activities.

4.1. History

The Census Bureau has a long history of working with administrative data. As far back as 1970, the Census Bureau has used administrative records to help enumerate group quarters, such as college dormitories and prisons – generally for situations where it is was not practical, possible, or safe to enumerate the inhabitants in person [18].

Prior to the 2000 Census, the Census Bureau began significant research on other uses for administrative records, including using them to enumerate non-responding households (instead of conducting costly in-person followup).

In 1991, Congress passed legislation [Public Law (P.L.) 102–135, the Decennial Census Improvement Act of 1991] requiring the National Academy of Sciences to study such things as the use of administrative records, sampling, and adjustment, and other ways to improve enumeration [19]. In December 1994, the Panel on Census Requirements in the Year 2000 and Beyond, an entity of the NAS’s National Research Council, issued a letter report to the Director of the Census Bureau stating “[it found] that it is not feasible to consider the use of a census based on administrative records for 2000.” [20] Another panel at the NRC focused on methodologies for the 2000 Census. They issued their final report in 1995, and also concluded “...that an administrative records census is not a feasible option for 2000.” [21]

About the same time, as part of the Census Bureau’s Task Force for Designing the 2000 Census, the Policy Committee of that Task Force was asked to study uses of administrative records for enumeration [22]. During the decade preceding 2000, the Census Bureau’s various advisory committees also studied and discussed uses of administrative records [23]. Also, as part of the research and testing before 2000, the Census Bureau studied a number of alternative census designs, including one that would make extensive use of administrative records for enumeration [24].

In 1992, the Census Bureau concluded that this would not be feasible for the 2000 Census [25]. The Census Bureau also recommended further research, including an experimental study as part of the 2000 Census. "AREX (Administrative Records Experiment) 2000 was the Census Bureau’s first attempt to use
administrative records as the foundation for a short-form decennial census. It examined the feasibility of conducting an administrative records census (ARC) as well as the use of administrative records as an ancillary method of data collection for the decennial census. AREX 2000 compared two methods of conducting an ARC. One relied solely on administrative records—such as birth and death records—and the other combined traditional enumeration methods with the use of administrative records. AREX 2000 also tested the potential uses of administrative records data for substitution processes and for other methods of defining and enumerating the nonresponse followup (NRFU) universe. One of the key findings from this test was the poor quality of these records with respect to data on race and Hispanic origin [26].

In 1999, the Census Bureau began development of the StARS system [27]. As stated in the current System of Records Notice (SORN): “The purpose of this system is to centralize and control the use of personally identifiable information by providing a secure repository that supports statistical operations through the removal of personal identifiers (Social Security Numbers and names), prior to delivery to other Census Bureau operating units. By combining current demographic and economic survey and census data with administrative record data from other agencies, and data procured from commercial sources on an as-needed basis, the Census Bureau will improve the quality and usefulness of its statistics and reduce the respondent burden associated with direct data collection efforts. The system will also be used to plan, evaluate, and enhance survey operations; improve questionnaire design and selected survey data products; and produce research and statistical products such as estimates of the demographic, social, and economic characteristics of the population.” [28]

The Census Bureau has established data sharing agreements with key stakeholders, such as the Internal Revenue Service and the Social Security Administration. These formal agreements are developed in conjunction with the data provider and vetted by all required offices. This includes a legal review by both agencies to ensure that the transfer is secure, and that the data will be protected within the Census Bureau network. In addition, the IRS conducts Safeguard Reviews on site at the Census Bureau at regular intervals to ensure that its data are being handled securely and in accordance with IRS laws, rules, and regulations. Each agreement spells out the data transfer methods. These are mutually agreed upon by both agencies. All transfers of data are fully encrypted using approved NIST FIPS 140-2 encryption. The exact systems and how the data are to be used are also spelled out in the agreement. In addition, an Information Security Agreement is prepared, reviewed and signed by the appropriate management and designated points of contact at each agency. Once at the Census Bureau, the data will reside on systems that are fully accredited according to NIST Special Publication 800-37r1 and assessed against NIST Special Publication 800-53r4 on a continuous basis.

The StARS system supports the entire Census Bureau, not just the decennial census program. For additional information about the development and management of StARS, see Prevost [29], Prevost and Leggieri [30], Blumerman, et al. [31], and Prevost and Obenski [32]. Also for further reading, a number of countries, particularly those with established population register systems, use administrative records extensively in conducting their censuses:

- United Nations description of population registers [33].
- United Nations Economic Commission for Europe (UNECE) reports on countries using population registers [34].
- Statistics Canada uses of administrative records [35].

5. Reengineering field operations for the 2020 Census

The goal of this innovation area is to use technology to efficiently and effectively manage the 2020 Census fieldwork, and as a result, reduce the staffing, infrastructure, and brick and mortar footprint required for the 2020 Census. The three main components of the reengineered field operations are streamlined office and staffing structure, increased use of technology, and increased management and staff productivity. In total, these design changes have the potential to save the Census Bureau an estimated $2.5 billion. The paper by Konicki and Adams [36] describes more details on some of the research the Census Bureau has done in this area.

5.1. History

The U.S. Census Bureau has a long history of innovation in the uses of automation, stretching back to the Hollerith Tabulator first used to tabulate results
for the 1880 Census [37]. It has used computers for data processing and tabulation since the early 1950s. In fact, in 1951 it was the first non-military agency to order a mainframe computer, UNIVAC I. [38] For the 1990 Census, the Census Bureau began using minicomputers in its temporary local/district offices (set up each decade to manage the large in-person data collection operations) [39]. Prior to the 1990 Census, it tested the feasibility of early portable computers to conduct address listing field operations [40]. After years of testing and development, the Census Bureau began using laptop computers for field data collection of the Current Population Survey in January 1994. Since that time, most surveys conducted by the Census Bureau have moved to laptop computers for field data collection. With its much larger workload (over 48 million Nonresponse Followup (NRFU) addresses in 2010, and over 500,000 temporary field interviewers) laptop computers would have been far too expensive to use for the decennial census. However, in 2009, the Census Bureau used small, handheld computers to conduct the nationwide Address Canvassing operation for the 2010 Census [41]. The Census Bureau also had planned to use these devices to conduct the NRFU operation in 2010. However, due to concerns about operational readiness and other factors, in 2008 the Census Bureau made a decision to cancel these plans and return to a paper-based NRFU for the 2010 Census [42].

According to the United Nations Economic Commission for Europe (UNECE) [43], several countries reported using handheld devices during the 2010 round of censuses. Brazil used them for field data collection [44], and Statistics Canada plans to use them for field data collection in their 2016 Census [45].

6. Overall 2020 Census research and testing plan

The 2020 Census has multiple decision points, milestones, and production dates that must be met to deliver the final apportionment and redistricting data by legal deadlines. Informing the decisions points are a series of tests [2]. More detailed information about each test is captured in formal research and test plan documents, in an integrated master schedule, facilitating the integration and coordination of activities across tests and operations, and in the 2020 Census Research and Testing Management Plan [46].

The tests conducted early in the decade (2012–2015) were designed to answer specific research questions (objectives) needed to make decisions on important aspects of the operational design for the four key innovation areas. Starting in 2016, the focus shifted to validating and refining the design by testing the interactions across operations and determining the final methodology for the operations. In addition, testing of production systems begins during this time frame, and will continue through 2018. An end-to-end test in 2018, with an April 1 Census Day, will be conducted to test the integration of all major operations and systems.

7. The 2020 Census operational plan – A new design for the 21st Century

Based on the innovations and research described above, the Census Bureau completed its 2020 Census Operational Plan by October 1, 2015 [2]. This design reflects a flexible approach that takes advantage of new technologies and data sources while minimizing risk.

8. Stakeholder and oversight activities

The Census Bureau is committed to maintaining an open and transparent process with stakeholders, including oversight groups. For example, a Monthly Status Report (MSR) is produced for the Census Bureau leadership and key staff, as well as key stakeholders [47]. The stakeholders include the U.S. Department of Commerce (the Census Bureau’s parent agency), auditors from the Office of Inspector General, the Office of Management and Budget, staff from authorizing and appropriations committees in Congress, and auditors from the Government Accountability Office.

A 2020 Census Program Management Review (PMR) is conducted quarterly and is open to the public. The purpose of the PMRs is to provide information and status (accomplishments, risks, challenges, near term focus, etc.) for the 2020 Census program. The PMRs are webcast live and archived on the Internet at census.gov [48] to allow stakeholders and interested parties to obtain current information about the 2020 Census program plans and progress.

Census Bureau leadership engage with other groups, including the Census Bureau’s formal advisory committees [49], and a panel of experts at the National Academy of Sciences [50]. In addition, leadership and staff actively participate at meetings to other groups, such as the American Statistical Association, the Population Association of America, the U.S. Conference of Mayors, and at various international forums.
The Census Bureau’s goal is a complete and accurate census in 2020 – counting everyone once, only once, and in the right place. It has been studying cost-saving design innovations for the last three years and is now shifting its focus to operationalizing the design and ensuring it will produce a quality census in 2020. The 2020 Decennial Census will look and operate quite differently than any previous census in the United States. Many of the innovations planned for 2020 are due to advances in data, methodologies and maturing technological capabilities. This is a census of firsts. The first census to prominently use the internet, the first census to use technology throughout the entire process, and the first census in many decades to make use of new data and methodologies. The release of the 2020 Census Operational Plan three years earlier than previous census cycles demonstrates the firm foundation for the planning of the 2020 Census.

**References**


[38] U.S. Census Bureau. UNIVAC I [Internet]. Available from: https://www.census.gov/history/www/innovations/technology/univac_1.html.


