Innovations in Census Taking for the United States in 2020

Burton H. Reist and Andrew Ciango
2020 Research and Planning Office, United States Census Bureau
Corresponding author: Burton H. Reist, email: Burton.H.Reist@Census.gov

Abstract

The 2010 Census produced exceedingly high quality data, but at the cost of $94 per housing unit. This was an increase of over 34 percent relative to 2000. Unless we make significant changes to the way we conduct the census, the cost will continue to rise at a similar rate in 2020. These escalating costs are unsustainable. The United States Census Bureau has embarked on a research and testing program to develop innovations in a wide range of operations designed to reduce costs while maintaining a high quality census. This includes leveraging the Internet and new communications strategies to improve self-response, determining the extent to which we can conduct a targeted address canvassing operation, and developing methods to use administrative records to enumerate a significant portion of the households that do not return a census questionnaire. We are also developing new methodologies that integrate and automate the conduct and management of field activities. Hiring and managing the staff needed to check addresses and conduct face-to-face interviews is the most expensive component of the census. While we know a large field staff will be needed regardless of the breakthroughs we make with administrative records and targeted address canvassing, the more we can cut back on this staff, enable them to work more efficiently, and streamline management, the more we can contain the costs of the census. This paper will provide an overview of the research and testing activities underway in each of these areas.

Introduction

The Census Bureau is committed to designing and conducting a 2020 Census that costs less per housing unit (on an inflation adjusted basis) than the 2010 Census, while maintaining high quality results. To achieve this goal, we must identify our cost drivers and implement innovative enumeration methods aimed at reducing these costs. Accordingly, we have embarked on an early research and testing program focused on major innovations to the design of the census oriented around the major cost drivers of the 2010 Census.

Background

The decennial census has been conducted since the birth of the United States. It provides the official population counts for determining the allocation to states of seats in the United States House of Representatives and for determining how the districts are defined for those seats.   The decennial census also provides comprehensive and useful demographic information about all people living in the United States, Puerto Rico, and

the associated Island Areas. Over $400 billion in federal funding is allocated annually using information from the decennial census. The provision of these data is mandated in the U.S. Constitution and Title 13 of the U.S. Code.

The cost of the 2010 Census was almost $94 per housing unit. This represented a 34 percent increase in the cost per housing unit over Census 2000 costs, which in turn represented a 76 percent increase over 1990 Census costs. If the Census Bureau makes no changes to the design of the decennial census, projected costs for the 2020 Census will increase at a similar rate

Analysis of the 2010 Census has determined that our costs were predominantly driven by five factors:

- the increasing diversity of the population;
- the demand for the Census Bureau to strive for improving accuracy over previous censuses;
- the lack of full public participation in the self-response phase of the census;
- the failure or challenges with linking major acquisitions, the schedule, and the budget; and
- substantial investments in major, national updating of the address frame just prior to enumeration (2009)

In order to control these costs for the 2020 Census, we initiated an aggressive and innovative research agenda. This agenda supports fundamental changes to program, business, operational, and technical processes to meet its vision and strategic goals. We plan to prove in innovations through research and iterative testing to engage in an evidence based decision-making process to support the major design decisions for the 2020 Census.

Research Agenda for the 2020 Census

The Census Bureau's early research agenda, conducted during the 2012-2015 timeframe, focuses on potential major innovation to the census design. The agenda concentrates on four major topic areas: optimizing self-response, administrative records, address list development, and improving field management and operations.

The rest of this paper will be devoted to the discussion of each of these individual topic areas. Each topic area discussion will explain how the research addresses the major cost drivers from the 2010 Census and will identify current and future research and testing activities for these topic areas. Our discussions will begin with optimizing self-response and end with improving field management operations.

Optimizing Self-Response

Lack of full public participation in the self-response phase of the census requires the hiring of a large field staff for Nonresponse Followup (NRFU) Operations. This is one of

the largest cost drivers for the census. If we can minimize the nonresponse universe, we can reduce the size of our field staff and infrastructure, thus reducing costs.

The traditional mail out/mail back procedures we have employed during previous censuses are no longer viable. The mail return rate has fallen from 87 percent in 1970 to 75.8 percent in 2010; see Treat (2004) and Letrouneau (2012). The increased diversity of the United State's population challenges our ability to rely extensively on the traditional mail procedures for self-response. Therefore, we must identify a mix of modes and strategies by demography and geography to increase self-response and reduce enumeration costs. During our research phase, we are exploring how we can leverage technology, variation in demographic/geographic response propensities, and new response modes to optimize self-response among a diverse population.

We are conducting research into the most effective modes to employ when contacting respondents. Towards this end, we are assembling an alternative contact frame consisting of email, landline, and cell phone contact information. We will analyze the alternate contact frame throughout a series of iterative field tests to assess the quality of the frame information. This research will determine what modes are most usable by demographic groups, geographic groups, and language groups.

We are also planning to optimize self-response by leveraging the internet. Research teams are exploring ways to drive respondents to the internet and take advantage of internet response processes to improve the data we collect. Plans are underway for iterative testing in this area and work is underway analyzing data from the American Community Survey and other surveys that use the internet.

Finally, we are researching and testing new communications strategies. Through a series of iterative tests conducted over the next couple of years, we plan to test several communication strategies and gauge the attitudes of the public with regard to our contact modes and response modes.

Administrative Records

The Census Bureau is committed to researching and testing the feasibility of using government and commercial sources of data in an effort to reduce the costs of the 2020 Census significantly. We are researching ways to use administrative records to reduce or eliminate some interviews of households that do not respond to the census and related field contacts while maintaining data quality. This would allow us to reduce the size of our field staff and infrastructure, thus lowering our costs.

In 2010, the Census Bureau initiated a study to match addresses and persons in administrative records to the 2010 Census and evaluate the quality and coverage of these administrative data. The 2010 Match Study indicated that the United States does not have the administrative records infrastructure suitable for a full administrative census, but administrative records could be used to supplement or replace the NRFU Operation and in support of other census operations; see Rastogi and O'Hara (2012). The match study

identified several key findings that support our decision to implement an early research agenda focused on the use of administrative data:

- Of the 2010 Census addresses, administrative records matched to 92.6 percent of them
- Of the 308.7 million persons in the 2010 Census, administrative records matched to 88.6 percent of them
- Of the 308.7 million person-address pairs in the 2010 Census, 70 percent matched to administrative records person-address pairs
- Of the 60.4 million person-address pairs in the 2010 NRFU operations 47.5 percent matched to administrative records; see Rastogi and O'Hara (2012)

We are continuing to research the applicability of administrative records to 2020 Census operations and addressing challenges identified in the match study, such as incomplete demographic information for particular groups. We are addressing the incomplete demographic information by profiling administrative records data sets to determine their fitness for use in 2020 Census operations.

We will begin field-testing the methodology for enumerating nonrespondents using a combination of administrative records and in-person visits during the test scheduled for October 2013. Iterative field tests are planned over the next couple of years that will allow us to refine our methodology for using administrative records to enumerate nonrespondents. The results of these field tests will allow us to prepare rough estimates of expected cost savings using this new methodology.

Address List Development

Our Master Address File (MAF) is the key control file for conducting the decennial census. It is a list of all the living quarters in the United States linked to their geographic location. Ensuring we have a complete and accurate MAF is a major effort and cost driver for the census. In previous censuses, we conducted a nation-wide, street-by-street canvassing effort to prepare this list, which required us to stand up an extensive field operation.

We are researching and testing methods to reduce the scope of the Address Canvassing operation for the 2020 Census. Our plan is to update the MAF continually throughout the decade, which will enable us to conduct a targeted address canvassing operation prior to the 2020 Census.

Our efforts to implement a targeted address canvassing operation prior to the 2020 Census have us undertaking research in two areas. First, we are researching a model to predict where coverage errors occur in the MAF. Secondly, we are assessing the Local Update of Census Addresses (LUCA) program to identify cost-effective changes, improve the quality of the MAF, and optimize the benefits derived by the Census Bureau and by participants (tribal, state, and local governments).

The work on the MAF coverage errors will allow us to create a model to predict where coverage errors occur in the MAF. To the extent that we can accurately model errors in the MAF, we will be able to target only specific areas where we need to update the address list prior to the 2020 Census.

We are also assessing the LUCA program and are conducting research to determine if it is possible to use available administrative data or "in house" means to validate these addresses. Many local government entities help us update our MAF through the LUCA program, which was established by Congress. If we can use existing records to validate some or all of the LUCA updates, we will not need to conduct costly fieldwork to verify the information on the ground.

Improving Field Management and Operations

We plan to improve the management of field enumerations by streamlining and automating operations and more efficiently planning and controlling interviewing activities. While we know an extensive field staff will be needed regardless of the breakthroughs we make with administrative records and targeted address canvassing, the more we can cut back on this staff, enable them to work more efficiently, streamline management, and increase productivity the more we can contain the costs of the 2020 Census.

We are researching methods to modernize Field Operations for the 2020 Census. This includes exploring innovative ways to reduce costs, and studying ways to develop integrated, Bureau-wide Information Technology (IT) infrastructure, capabilities and cost-effective IT technical solutions, which can lead to reduced costs by allowing the Census Bureau to move towards a more flexible and scalable infrastructure.

One way we plan to reach our goal of reducing field costs is by investigating ways we can increase efficiency through further automating field activities. We believe that by developing better IT systems for controlling work in the field and leveraging mobile applications on hand-held devices for managing and conducting field operations, we can increase our efficiency in the field. We are currently in the process of researching and developing tools for automated time and expense reporting and route planning for each enumerator to use turn-by-turn directions.

Another way we plan reduce field costs is through researching and testing adaptive design methods to supplant the fixed design we have traditionally used for nonresponding households. The adaptive design framework makes use of auxiliary frame data and paradata to continually update case assignments and develop a contact strategy tailored to each household with the goal of reducing and improving person follow-up activities.

We plan to begin field-testing our adaptive design approach as part of the October 2013 test. This will be the first use of an adaptive design approach in a quasi-census environment. Personal interviewers in the adaptive design contact strategy component will be told, on a daily basis, which cases are priorities for contact that day and when to perform proxy interviews, as determined by response likelihood

and cost models. The results of this test will inform the large-scale implementation of the adaptive contact strategy component planned for the 2014 Census Test.

Conclusion

As this research agenda demonstrates, the Census Bureau is committed to identifying and implementing innovative solutions to the challenge of reducing cost and maintaining quality for the 2020 Census. We have identified the major cost drivers from the 2010 Census and have implemented a research and testing agenda to address those issues early in the 2020 Census lifecycle.

In addition to the above-mentioned research, the Census Bureau is addressing the rising costs related to the challenges of linking major acquisitions, the schedule, and the budget by investing in a program management framework to undergird the 2020 Census Program. This framework is based on industry best practices, lessons learned from the 2010 Census, and alignment with initiatives at the agency level. It provides the 2020 Program with the strategies, processes, and tools recognized by the disciplines of Budget Management, Schedule Management, and Sourcing and Acquisition Management. The Census Bureau also will better define requirements for all systems through a robust systems engineering and integration discipline that already is under development.

Our original goal was to make major design decisions, based on this research, by the end of fiscal year 2014. However, because of a reduction in our fiscal year 2013 budget, our opportunities to incorporate innovation and test cost savings methods have been reduced or delayed. This means we are losing opportunities to iteratively test, refine, and re-test before we make decisions regarding the operational design. This increases the risk that we will not have sufficient time for designing, developing, and testing systems necessary to implement the census, particularly new cost saving innovations that have not been used in previous censuses. In an effort to mitigate this risk, we revised our original testing strategy and delayed our operational design decisions from the end of fiscal year 2014 to the end of fiscal year 2015.

References

Letrouneau, E. (2012) "Mail Response/Return Rates Assessment," United States Census Bureau, 14.

Rastogi, S. and O'Hara, A. (2012) "2010 Census Match Study," United States Census Bureau, ix -xii.

Treat, J. (2004) "Census 2000 Testing, Experimentation, and Evaluation Program: Response Rates and Behavior Analysis," United States Census Bureau, 17.