



U.S. CENSUS BUREAU

2020 Census: 2016 Census Test Indicates the Current Life-Cycle Cost Estimate is Incomplete and Underestimates Nonresponse Followup Costs

FINAL REPORT NO. OIG-17-020-I
MARCH 16, 2017

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

FOR PUBLIC RELEASE





March 16, 2017

MEMORANDUM FOR: John H. Thompson
Director
U.S. Census Bureau

FROM: 
Carol N. Rice
Assistant Inspector General for Economic
and Statistical Program Assessment

SUBJECT: *2020 Census: 2016 Census Test Indicates the Current Life-Cycle Cost Estimate is Incomplete and Underestimates Nonresponse Followup Costs*
Final Report No. OIG-17-020-1

Attached for your review is our final report on the evaluation of the Census Bureau's 2020 decennial program preparation and planning efforts. We initiated our evaluation of the 2016 Census Test to review (1) the effectiveness of the new management structure and (2) the effectiveness of the operational control system (OCS) to support nonresponse followup (NRFU) operations. Our first objective was to determine whether the Bureau documented its decision to utilize the enumerator-to-supervisor ratios selected for the 2016 test. Our second objective was to determine whether the Bureau designed the test to (a) assess the effectiveness of the OCS in supporting supervisors during NRFU operations by comparing 2016 test results to results of previous tests and (b) determine the feasibility and effectiveness of a higher enumerator-to-supervisor ratio compared to the 2010 Census.

We found the following:

- More than 10 million potential NRFU contact attempts are unaccounted for in the life-cycle cost estimate.
- Limitations to the design and methodology of the 2016 Census Test hinder the Bureau's ability to answer research questions.
- Improvements could increase the effectiveness of the OCS.
- Field staff were not adequately trained to complete proxy interviews and receive paper questionnaires.

In addition, we identified that the 2016 Census Test Study Plan included inaccurate statements and assumptions about past research.

We have summarized the Bureau's response to our draft report, as well as included it as appendix B. The final report will be posted on OIG's website pursuant to sections 4 and 8M of the Inspector General Act of 1978, as amended (5 U.S.C. App., §§ 4 & 8M).

In accordance with Departmental Order 213-5, please submit to us—within 60 calendar days of the date of this memorandum—an action plan that responds to the recommendations of this report.

We appreciate the cooperation and courtesies extended to us by your staff during our review. If you have any questions or concerns about this report, please contact me at (202) 482-6020 or Terry Storms, Supervisory Auditor, at (202) 482-0055.

Attachment

cc: Laura Furgione, Chief of the Office of Strategic Planning, Innovation and Collaboration
Lisa M. Blumerman, Associate Director for Decennial Census Programs, Census Bureau
Colleen T. Holzbach, Program Manager for Oversight Engagement, Census Bureau
Pamela Moulder, Senior Program Analyst, Economic and Statistics Administration
Corey J. Kane, Audit Liaison, Census Bureau



Report in Brief

March 16, 2017

Background

The Census Bureau (the Bureau) recognizes that fundamental changes to the design, implementation, and management of the 2020 Census must occur in order to conduct the next decennial census at a lower cost (per housing unit and adjusted for inflation) than the 2010 Census. The 2010 Census cost approximately \$13 billion. The Bureau estimates that if it were to conduct the 2020 Census just as it conducted the 2010 Census, then the next decennial census would cost \$17.8 billion. By implementing a number of innovations, the Bureau expects that the 2020 Census will cost \$12.5 billion, avoiding \$5.3 billion of the estimated cost.

Why We Did This Review

We initiated our evaluation of the 2016 Census Test to review (1) the effectiveness of the new management structure and (2) the effectiveness of the OCS to support NRFU operations. Our first objective was to determine whether the Bureau documented its decision to utilize the enumerator-to-supervisor ratios selected for the 2016 test. Our second objective was to determine whether the Bureau designed the test to (a) assess the effectiveness of the OCS in supporting supervisors during NRFU operations by comparing 2016 test results to results of previous tests; and (b) determine the feasibility and effectiveness of a higher enumerator-to-supervisor ratio compared to the 2010 Census.

U.S. CENSUS BUREAU

2020 Census: 2016 Census Test Indicates the Current Life-Cycle Cost Estimate is Incomplete and Underestimates Nonresponse Followup Costs

OIG-17-020-I

WHAT WE FOUND

During our assessment of cost data and the cost estimate—just as we found during prior tests—we identified problems with the Bureau’s method for calculating potential cost avoidance during the 2020 Census. The current cost estimate contains assumptions that underestimate nonresponse followup (NRFU) costs.

Regarding our second objective we found that the Bureau cannot (1) determine the feasibility and effectiveness of a higher enumerator-to-supervisor ratio because flaws to the design and methodology of the test hinder the Bureau’s ability to answer research questions, or (2) assess the operational control system’s (OCS’s) capacity to effectively and efficiently manage NRFU operations.

Finally, we identified training limitations, which potentially impact the quality and protection of household data collected by enumerators. Also, we found that the 2016 Census Test Study Plan included inaccurate statements and assumptions about past research.

The 2020 Census life-cycle cost estimate assumes: (1) housing unit visits are limited to six attempts; and (2) all housing units are enumerated after the maximum number of allowed contact attempts. We found that this is not the case and if not corrected, NRFU costs will be underestimated. According to the Government Accountability Office, a credible cost estimate should include an independent review of the estimate and recognize and document excluded costs.

WHAT WE RECOMMEND

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

1. Ensure that the 2020 Census life-cycle cost estimate accurately reflects all relevant cost factors and excluded costs are documented.
2. Designate appropriate personnel to independently verify that tests are properly designed to answer research questions.
3. Work with the Office of General Counsel to develop enumerator scheduling guidance; and implement the most efficient NRFU enumeration scheduling practices.
4. Implement internal controls in the new operational control system that (a) prevent supervisors from ignoring alerts and inform managers that alerts were not responded to in a timely manner; and (b) provide supervisors with sufficient detail to resolve alerts.
5. Revise training to ensure field staff are adequately prepared to conduct proxy interviews and securely transmit paper questionnaires for processing.

Contents

Introduction	1
Objectives, Findings, and Recommendations	3
I. More Than 10 Million Potential NRFU Contact Attempts Are Unaccounted for in the Life-Cycle Cost Estimate.....	4
Increased rate of unresolved NRFU cases may require additional resources.....	5
II. Limitations to the Design and Methodology of the 2016 Census Test Hinder the Bureau’s Ability to Answer Research Questions	6
III. Improvements Could Increase the Effectiveness of the OCS	7
A. The Bureau did not require enumerators to work when household respondents were most likely to be home.....	7
B. Supervisors did not consistently respond to alerts.....	8
C. Alerts did not differentiate between production and quality control cases	9
IV. Field Staff Were Not Adequately Trained to Complete Proxy Interviews and Receive Paper Questionnaires	9
Recommendations	10
Other Matters	11
Summary of Agency Response and OIG Comments	12
Appendix A: Objectives, Scope, and Methodology	13
Appendix B: Agency Response	15

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

Introduction

The Census Bureau (the Bureau) recognizes that fundamental changes to the design, implementation, and management of the 2020 Census must occur in order to conduct the next decennial census at a lower cost (per housing unit and adjusted for inflation) than the 2010 Census. The 2010 Census cost approximately \$13 billion. The Bureau estimates that if it were to conduct the 2020 Census just as it conducted the 2010 Census, then the next decennial census would cost \$17.8 billion. By implementing a number of innovations, the Bureau expects that the 2020 Census will cost \$12.5 billion, avoiding \$5.3 billion of the estimated cost. These innovations involve four key areas:

1. new methodologies to conduct address canvassing;¹
2. options to increase household self-response;²
3. the use of administrative records to reduce the nonresponse followup (NRFU) workload; and
4. reengineering field operations through the use of technology replacing intensive paper processes (such as enumeration, payroll, and training), realigning staff, and reducing the number of field offices.

The 2016 Census Test was intended to allow the Bureau to study new methods that are under consideration for the 2020 Census NRFU operation.³ Specifically, building on testing conducted in 2014 and 2015, the Bureau intended to use the 2016 test to refine technologies and methods for assigning housing units to enumerators to visit and collect household information. Design changes associated with the NRFU operation are expected to account for \$2.5 billion of the Bureau's estimated cost avoidance.⁴ The 2016 test was conducted at two locations: Los Angeles County (Los Angeles), California, and Harris County (Houston), Texas. The combined geographic areas were anticipated to include approximately 450,000 housing units to be contacted in an initial self-response phase, followed by a NRFU phase of approximately 120,000 non-responding housing units.

¹ Address canvassing is the process by which the Bureau validates, corrects, or deletes existing addresses, adds missing addresses, and adds or corrects locations of specific addresses before decennial census enumeration activities.

² One of the Bureau's general objectives is to increase self-response by making it easier to respond to the decennial and household surveys. This includes the use of Internet, innovative contact strategies, and new methods of contacting the public, such as telephone and e-mail.

³ NRFU is an operation that sends enumerators to housing units that the Bureau did not receive a response by Internet or mail.

⁴ All values conveyed in 2020 constant dollars.

The Bureau designed the 2016 test to investigate a number of 2020 Census design options (see graphic). Most relevant to the objectives of our evaluation were the Bureau's efforts to assess two different enumerator-to-supervisor staffing ratios (20:1 in Los Angeles; and 30:1 in Houston). During the 2010 Census, supervisors in the field managed approximately eight enumerators. With the use of automated systems and other innovative decennial design improvements, the Bureau believes that supervisors can manage more than eight enumerators.

During this test, two systems⁵ were used to perform certain activities, such as creating the NRFU universe, creating enumerator assignments, and tracking and reporting on enumerator performance. These two operational control systems (OCSs) were used in prior tests; the 2016 test was another opportunity for the Bureau to refine the system requirements that will inform the actual system used during the 2020 Census. Specific to our evaluation objectives, the 2016 test was an opportunity for the Bureau to further assess planned automation, such as optimized enumerator assignments and methods to track and report on enumerator performance. Each optimized daily assignment included a list of cases and a predetermined sequence in which the cases were to be attempted based on the most efficient route between each address. The OCS also generates real-time data, including system-generated alerts that indicate to a supervisor that an enumerator's performance varied from what was expected and requires attention and a response.

Enumerators used a mobile application called COMPASS,⁶ to submit work availability, time, and expense data; receive assignments; record NRFU contact attempts; and collect household data. The COMPASS application was loaded onto smartphones, which the Bureau provided to enumerators for the duration of the test.

2020 Census Design Components Tested During the 2016 Census Test

- self-response options
- use of existing government and commercial information to reduce the NRFU workload
- partnership and outreach efforts with historically hard-to-count populations
- different enumerator-to-supervisor staffing ratios
- technologies and methods for assigning, managing, and collecting enumerator fieldwork

Source: U.S. Census Bureau

⁵ The first system is the multi-mode OCS that was designed to create and manage the nonresponse workload for enumerator-respondent contact. The second system is the enhanced OCS, with near real-time monitoring and decision-making capabilities during field activities. For the duration of the report, when we refer to the OCS, we are referring to the enhanced OCS.

⁶ COMPASS (Census Operations Mobile Platform for Adaptive Service Solutions) collects respondent data and associated enumerator information on mobile devices.

Objectives, Findings, and Recommendations

We initiated our evaluation of the 2016 Census Test to review (1) the effectiveness of the new management structure and (2) the effectiveness of the OCS to support NRFU operations. Our first objective was to determine whether the Bureau documented its decision to utilize the enumerator-to-supervisor ratios selected for the 2016 test. Our second objective was to determine whether the Bureau designed the test to (a) assess the effectiveness of the OCS in supporting supervisors during NRFU operations by comparing 2016 test results to results of previous tests and (b) determine the feasibility and effectiveness of a higher enumerator-to-supervisor ratio compared to the 2010 Census. To accomplish our objectives, we interviewed officials at Census Bureau headquarters, reviewed documentation, and observed enumerators, supervisors, and other staff in the field. For further discussion regarding our objectives, scope, and methodology, see appendix A.

Although not an original objective of this evaluation, after we identified internal control weaknesses that could affect the Bureau's ability to achieve cost savings, we assessed (1) the Bureau's ability to collect cost data during the 2016 test and (2) whether the cost estimate accounted for cases that were still unresolved after the maximum number of contact attempts allowed. Cases that remain unresolved following the maximum number of NRFU contact attempts negatively impact the quality of data collected for that area since other, less-reliable means of data collection such as imputation⁷ must be used to enumerate those housing units. One goal of the 2016 test was to validate assumptions associated with key NRFU cost parameters included in the 2020 Census life-cycle cost estimate. During our assessment of cost data and the cost estimate—just as we found during prior tests⁸—we identified problems with the Bureau's method for calculating potential cost avoidance during the 2020 Census. The current cost estimate⁹ contains assumptions that underestimate NRFU costs (see finding I).

With respect to our first objective, evidence indicates that the Bureau can properly support its enumerator-to-supervisor ratio decisions. Our review of the Bureau's method for selecting the optimal ratios for the test indicate that that the Bureau received two contracted reports from independent sources to inform its decision. Based on this information, the Bureau decided that each test location would employ only one of the recommended ratios (an average of 20 enumerators for every supervisor in Los Angeles, and an average of 30 enumerators for every supervisor in Houston).

Regarding our second objective, we found that the Bureau cannot (1) determine the feasibility and effectiveness of a higher enumerator-to-supervisor ratio because flaws to the design and

⁷ Imputation assigns values for missing or inconsistent information by relying on available information from other persons or households with similar characteristics located within a small geographic area.

⁸ (1) U.S. Department of Commerce, Office of Inspector General, June 7, 2016. *2020 Census: The Bureau Has Not Reported Test Results and Executed an Inadequately Designed 2015 Test*, OIG-16-032-A. Washington, DC: OIG. (2) U.S. Department of Commerce, Office of Inspector General, September 30, 2015. *2020 Census: The 2014 Census Test Misses an Opportunity to Validate Cost Estimates and Establish Benchmarks for Progress*, OIG-15-044-A. Washington, DC: OIG.

⁹ The Bureau's most current life-cycle cost estimate is dated October 1, 2015.

methodology of the test hinder the Bureau's ability to answer research questions (see finding II), or (2) assess the OCS's capacity to effectively and efficiently manage NRFU operations (see finding III).

Finally, we identified training limitations, which potentially impact the quality and protection of household data collected by enumerators (see finding IV). Also, we found that the 2016 Census Test Study Plan included inaccurate statements and assumptions about past research (see Other Matters).

The 2020 Census life-cycle cost estimate assumes: (1) housing unit visits are limited to six attempts; and (2) all housing units are enumerated after the maximum number of allowed contact attempts. We found that this is not the case and if not corrected, NRFU costs will be underestimated. According to the Government Accountability Office, a credible cost estimate should include an independent review of the estimate and recognize and document excluded costs.¹⁰

I. More Than 10 Million Potential NRFU Contact Attempts Are Unaccounted for in the Life-Cycle Cost Estimate

If a member of a household does not respond to a decennial census questionnaire, or through another mode of self-response, then that housing unit is included in the NRFU universe, and will be assigned to an enumerator who attempts to collect household data during a personal visit. As part of the 2016 Census Test, the Bureau planned to "evaluate the impacts on cost and quality of a NRFU contact strategy that allows for a maximum of six contact attempts." However, during the 2016 test, by analyzing contact attempts data provided by the Bureau, we found that 10 percent of the housing units in the NRFU workload received more than 6 attempts, resulting in an additional 29,411 visits. We reported similar results in our 2015 test report. These additional attempts occurred during the 2016 test because the OCS was programmed to limit contact attempts to 6 days rather than 6 attempts in total. For example, a housing unit could receive 1 contact attempt on the first day and 10 contact attempts on the second day and still be eligible for more attempts on 4 more days. Our analysis identified 1 enumerator who contacted a housing unit 11 times in 1 day.

According to the program managers we spoke to, they were unaware that the OCS did not remove NRFU cases from the workload after the sixth contact attempt. The almost 30,000 additional housing unit visits could equal more than 11 million¹¹ attempts during the actual decennial. The Bureau must align its contact strategy (6 days versus 6 attempts) with the cost estimate, otherwise the estimate will understate costs—in terms of enumerator salary and mileage—because of the additional attempts.

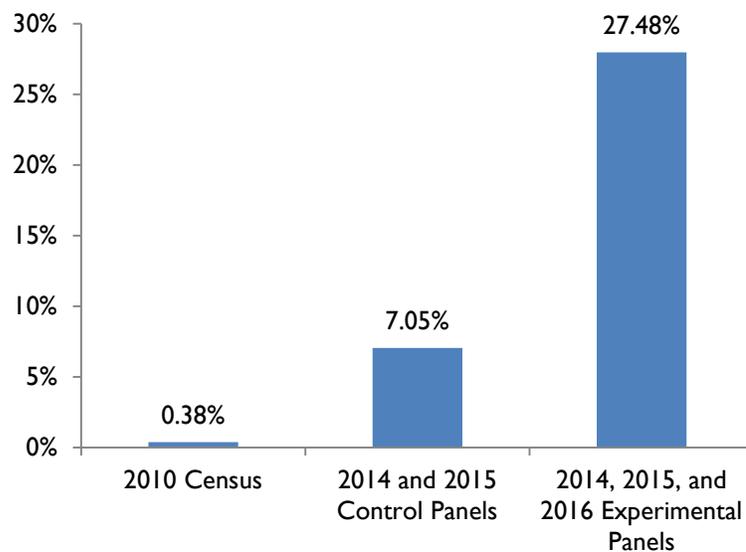
¹⁰ Government Accountability Office, March 2009. *GAO Cost Estimating and Assessment Guide: Best Practices for Developing and Managing Capital Program Costs*, GAO-09-3SP. Washington, DC: GAO.

¹¹ During the 2016 test, there were 29,411 attempts over the maximum for a total workload of 143,786 cases, or 0.205 attempts over maximum per case. 2020 NRFU Assumptions and Costs show a total NRFU workload of 56,344,671. Multiplying the expected 2020 workload by the attempts over maximum per case during the 2016 test equals 11,550,658 (56,344,671 * 0.205%).

Increased rate of unresolved NRFU cases may require additional resources

As previously stated, the current cost estimate assumes that all NRFU cases will be successfully enumerated by the end of the NRFU operation. However, during the 2010 Census, the Bureau ceased NRFU operations on 521,947 housing units (0.38 percent of U.S. housing units). These housing units were then enumerated through other, less desirable methods, such as imputation.¹² Although less than one percent of U.S. housing units were unresolved following the 2010 Census NRFU, tests of new innovations—that the Bureau expects to improve NRFU efficiency during the 2020 Census—conducted in 2014, 2015, and 2016 resulted in much higher rates of unresolved cases (see figure 1).

Figure 1. 2010 Census NRFU Housing Unit Unresolved Rates Compared to the 2014, 2015, and 2016 Census Tests



Source: U.S. Census Bureau

^a During the 2014 and 2015 Census Tests, the Bureau made use of control panels, which conducted NRFU much the same as it was conducted during the 2010 Census in order to measure the effect on NRFU of new innovative techniques—used by experimental panels—that it is considering implementing during the 2020 Census. The 2016 test did not use a control panel.

The low rate of unresolved cases observed during the 2010 Census was likely due, in part, to strategies intended to increase response rate that were not implemented during tests conducted in 2014, 2015, and 2016 (e.g., nationwide publicity, the Census partnership program, paid advertising). The rate of unresolved cases observed during recent tests is notably higher than the 2010 Census rate. If new innovative NRFU

¹² Count imputation assigns housing unit status or household size to records that are missing that information. The Bureau used count imputation to enumerate unresolved housing units during the 2010 Census; the Bureau plans to use administrative records and other outside data sources to enumerate nonresponding housing units during the 2020 Census.

procedures result in an increased unresolved rate, the Bureau will have to expend additional resources—which are not currently accounted for in the cost estimate—to fulfill the Constitutional requirement to count the population.

We believe inaccurate cost estimating is occurring in part because of poor communication between those responsible for developing the cost estimate and those responsible for testing NRFU strategies. The Bureau must take steps to ensure that the 2020 Census life-cycle cost estimate accurately reflects all relevant cost factors and documents costs excluded from the estimate.

II. Limitations to the Design and Methodology of the 2016 Census Test Hinder the Bureau's Ability to Answer Research Questions

A primary goal of the 2016 test was to further investigate the enumerator-to-supervisor staffing ratio in order to determine how many enumerators each supervisor can effectively manage during NRFU operations. However, similar to the 2015 Census Test, the Bureau did not properly control all of the other factors that could have influenced NRFU performance during the test; therefore, the Bureau is unable to isolate the effect of each ratio on NRFU performance. For example, the Bureau only used one ratio at each location (30:1 in Houston and 20:1 in Los Angeles). Other factors that were not controlled, but likely had an effect on performance, included:

1. devices—enumerators in Houston used iPhones® while enumerators in Los Angeles used Android™ phones;
2. the use of unique closeout procedures to try to resolve cases in Los Angeles but not in Houston; and
3. staff in Houston were provided instructions to ignore the sequence of housing unit visits, but those same instructions were not provided in Los Angeles.

To fairly test and measure the effect on NRFU performance of each enumerator-to-supervisor ratio, the Bureau should have controlled for the effect of other factors—such as location and mobile devices, along with any other factors that affect NRFU performance—by randomly assigning enumerators and supervisors to various groups (e.g., 20:1 ratio and iPhone, and 30:1 ratio and iPhone) at both locations. Finally, eliminating inconsistencies between groups (e.g., unique Los Angeles closeout procedures) during the test would help ensure valid test results.

Due to 2016 test design and methodology limitations, the Bureau will not be able to adequately answer some of its basic research questions, such as:

- What are the cost and quality tradeoffs associated with each staffing scenario?¹³
- Do supervisors who manage more enumerators generally cost more?¹⁴

¹³ U.S. Census Bureau, March 15, 2016. *2020 Census: 2016 Census Test Plan*, Version 3.0. Washington, DC: Census, p. 32.

¹⁴ U.S. Census Bureau, July 29, 2016. *2016 Census Test Study Plan: Nonresponse Followup (NRFU)*. Washington, DC: Census, p. 9.

- Do supervisors who manage more enumerators have a higher or lower completion rate?¹⁵

The Bureau's inability to isolate the effect of each enumerator-to-supervisor ratio on NRFU performance (e.g., measures of cost, quality, and completion rate) means that the Bureau is unable to determine the optimal enumerator-to-supervisor staffing ratio.

III. Improvements Could Increase the Effectiveness of the OCS

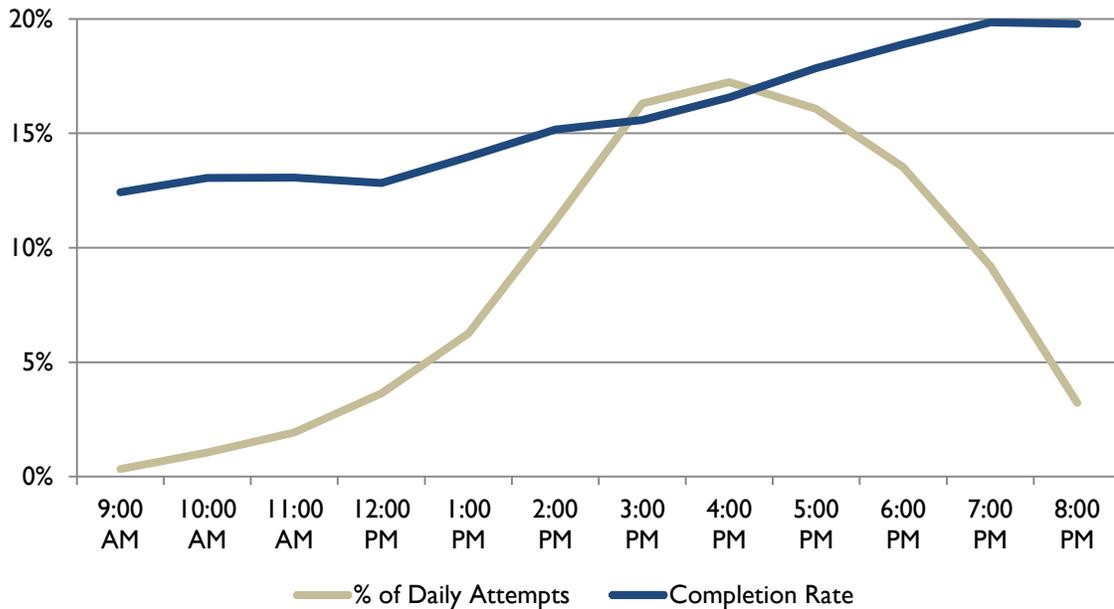
The Bureau is replacing the OCS that was used during the 2015 and 2016 Census Tests through a contract with an outside vendor. In addition to providing near real-time information for monitoring NRFU activities and operational decision-making, as well as creating optimized daily assignments for enumerators, the OCS includes an alerts function that automatically generates alerts, or messages, to supervisors if an enumerator's performance varies from what is expected. However, we identified three issues that could be addressed in future iterations of the OCS: (1) the Bureau did not require enumerators to work when household respondents were most likely to be home, and nothing prevented an enumerator from working outside of the hours included in his or her daily assignment; (2) supervisors did not consistently respond to alerts; and (3) alerts did not differentiate between production and quality control cases.

A. *The Bureau did not require enumerators to work when household respondents were most likely to be home*

The Bureau conducted research to analyze the optimal time of day to make contact with a household member at a nonresponding unit. Using paradata¹⁶ from the American Community Survey (ACS), the Bureau created simulation models to predict the probability of contacting someone at home for each daytime hour. Although the Bureau had data that showed enumerators had a higher probability to complete an interview between 2:00 p.m. and 7:00 p.m., the OCS did not require enumerators to work during these hours. We analyzed Bureau data and found that the number of contact attempts by enumerators declined between 4:00 p.m. and 8:00 p.m., compared to other times indicating that fewer enumerators worked during hours with the highest completion rate (see figure 2).

¹⁵ Ibid.

¹⁶ Paradata are data about the process by which survey information is collected. Due to the emergence of computer assisted surveys, paradata are now able to be captured through key-strokes, time stamps of movement, and navigation patterns throughout the questionnaire.

Figure 2. Contact Attempt Rate vs. Completion Rate for Weekdays

Source: OIG analysis of Census Bureau data

Staff who programmed the OCS believed that the Bureau could not legally restrict enumerator working hours; however, our review of the Bureau's internal discussion with the Department's Office of General Counsel indicated that the issue of restriction of an enumerators' work week was only discussed in the context of the requirement of premium pay for work on Sundays and did not contain any general prohibition against restricting an enumerator's hours. As a result of this misunderstanding, the Bureau did not develop controls to assign work during evening peak hours or establish controls to prevent an enumerator from working his or her cases outside of their scheduled shift. For example, if an enumerator was scheduled to begin work at 3:00 p.m., the OCS did not prevent them from starting earlier. In addition, NRFU contact attempts made when household respondents were not as likely to be home may increase the mileage and salary costs of enumeration during the 2016 test.

B. Supervisors did not consistently respond to alerts

Supervisors failed to act on 2,776 of the 12,047 (23 percent) alerts generated during the 2016 test because the OCS did not require supervisors to respond to alerts. Alerts help reduce the risk of both cost overruns in the form of excessive pay or mileage claims, and data falsification. For example, a supervisor receives an alert if an enumerator claims more miles or hours worked than expected based on the given assignment. Alerts are also generated when enumerators submit interviews that take very little time (i.e., less than 2 minutes), or occur too far from a housing unit (i.e., 3,000 meters or more), actions indicating potential fraud. However, Field Managers of Operations—who managed NRFU operations—were unable to see whether supervisors had alerts that were about to expire or had already expired. Managers for both the Houston and Los

Angeles test locations complained that they lacked tools to review the status of alerts. In addition to cost implications, data quality issues (e.g., data falsification) may not be detected or prevented if supervisors are not researching and resolving alerts.

During the 2015 test, we reported that the OCS did not require supervisors to act upon alerts and, as a result, 15 percent of the 3,329 alerts generated expired without any supervisory action. In response to this audit finding, the Bureau stated that it would use the knowledge, experiences, and lessons learned during the 2015 test to implement adjustments to the OCS for the 2016 Census Test. The 2016 test's goal was to "refine[e] and expand alerting capabilities."¹⁷ Based on the current review, there is still no control in place to ensure that supervisors review and resolve alerts.

C. Alerts did not differentiate between production and quality control cases

In order to resolve alerts and identify unusual activity, supervisors must be able to identify the type of case being worked. However, when reviewing cases during the 2016 test, supervisors were unable to determine whether enumerators were working on production cases or quality control cases,¹⁸ complicating their ability to respond to certain alerts. For example, a short interview alert is generated if an enumerator completes an interview in less than 2 minutes; however, quality control interviews can be brief. For the short interview alert to be effective, supervisors must know whether the alert was generated for a production interview or quality control interview. This functionality was not included in the 2016 test.

Alerts are only effective if supervisors use and adequately resolve them. If there is no internal control in place to make supervisors take action on alerts, some alerts may be actively ignored or expire before a supervisor takes action, thereby reducing the impact that alerts can have on cost containment and data quality. Likewise, alerts are not meaningful if a supervisor does not have sufficient detail to resolve the alert.

IV. Field Staff Were Not Adequately Trained to Complete Proxy Interviews and Receive Paper Questionnaires

During site visits at both Houston and Los Angeles, we noted that enumerators may not have been trained to accurately obtain information from proxy respondents¹⁹ or turn in paper questionnaires received from household respondents. As a result, household information received during the 2016 test may be incomplete, of questionable quality, or lost.

Specifically, we noted during a site visit in Los Angeles an instance where an enumerator used himself as a proxy to designate a housing unit as vacant, despite not verifying if that

¹⁷ Census, *2020 Census: 2016 Census Test Plan*, March 15, 2016, p. 6.

¹⁸ Following enumeration, some NRFU cases are subject to a quality control interview, in which a different enumerator visits the housing unit to ensure that an interview took place.

¹⁹ A proxy respondent is an individual who is not a member of the household being enumerated, such as a neighbor or some other knowledgeable person.

was accurate as of April 1, 2016, the test's official Census Day. The 2016 test included a total of 24,160 proxy attempts in both Los Angeles and Houston. A housing unit becomes proxy eligible on the third attempt to enumerate the case. Of the 24,160 attempts, 5,415 (22.4 percent) cases were completed by proxy respondent. By nature, proxy interviews may be less accurate than interviewing a household member. However, if enumerators are not trained to adhere to guidelines when obtaining proxies, the information obtained may be inaccurate and unreliable. Given the number of potential proxy interviews during the actual decennial census,²⁰ the Bureau should ensure that enumerators are adequately trained to collect accurate household data through proxy interviews.

Procedures for handling paper questionnaires were not included in enumerator or supervisor training. We identified 217 instances where enumerators received paper questionnaires from household respondents. During our visit to the Los Angeles regional office, we noted that the Field Manager of Operations had three paper questionnaires that were given to enumerators by household respondents. The Field Manager of Operations did not know how to submit the paper questionnaires to the National Processing Center, even though a procedure was in the Field Manager of Operations training manual. When we asked the Bureau to verify that the paper questionnaires were received, the Bureau was unable to tell us how many of the 217 questionnaires were processed.

Inadequate enumerator and supervisor training can result in incomplete and unreliable household data. In order for the Bureau to have the most accurate and reliable data, field staff must be properly trained to conduct proxy interviews and process paper questionnaires. Additionally, lack of control over paper questionnaires that contain household data potentially compromises the protection of Title 13 data.

Recommendations

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

1. Ensure that the 2020 Census life-cycle cost estimate accurately reflects all relevant cost factors and excluded costs are documented.
2. Designate appropriate personnel to independently verify that tests are properly designed to answer research questions.
3. Work with the Office of General Counsel to develop enumerator scheduling guidance; and implement the most efficient NRFU enumeration scheduling practices.
4. Implement internal controls in the new operational control system that (a) prevent supervisors from ignoring alerts and inform managers that alerts were not responded to in a timely manner; and (b) provide supervisors with sufficient detail to resolve alerts.
5. Revise training to ensure field staff are adequately prepared to conduct proxy interviews and securely transmit paper questionnaires for processing.

²⁰ During the 2010 Census more housing units (24.4 million) were enumerated via proxy than were enumerated via interview with a household member (22.2) million.

Other Matters

The Bureau developed the 2016 Census Test Study Plan to evaluate the NRFU technologies and methods used during the 2016 Test to help refine those technologies and methods for future operations, as well as ensure that the 2020 Census NRFU design is efficient and effective. However, we found that this study plan included inaccurate statements and assumptions about past research (i.e., the 2014 and 2015 Census Tests). The Bureau admitted that certain statements included in the study plan were overstated or misstated because they were accepted as statements of what was actually done during past tests, when in fact the statements were simply conclusions reached by staff who contributed to the study plan and that were not supported by previous test results. For example, the study plan stated that the 2014 and 2015 tests examined the enumerator-to-supervisor ratio and determined that the 2015 proportion of 15:1 was “sufficient.” However, the 2014 and 2015 tests did not test—or measure differences between—various ratios; those tests only utilized one ratio, so there were no conclusions to be made. The Bureau does not intend to revise the study plan to remove or update these inaccurate statements and assumptions.

Summary of Agency Response and OIG Comments

On February 24, 2017, we received the Census Bureau's comments on the draft report, which we have included as appendix B of this final report. The Bureau agreed with all five recommendations.

In response to finding I, the Bureau concurred with our recommendation but questioned our assumptions and extrapolations. Specifically, the Bureau stated "the life-cycle cost model already includes the rates of non-interview after six NRFU contact attempts that we assume for a full and accurate count in the 2020 Census. It also includes funding for various situations that might require more field visits than expected." However, during our review, the Bureau did not provide evidence to support these statements. Rather, discussions with Bureau management and our analysis indicate that the current life-cycle cost estimate assumes that all non-responding households will be resolved after six attempts.

Additionally, the Bureau disagreed with our attempt to extrapolate test results to the 2020 Census. We recognize that Houston and Los Angeles—the sites selected by the Bureau for the 2016 test—are not statistically representative of the entire United States. Thus, estimates based on the 2016 test results are not generalizable nationwide. However, the results still illustrate potential cost challenges for the 2020 Census that the Bureau should consider. For example, while the Bureau stated "there is no evidence that nearly all cases cannot be resolved within six NRFU contact attempts," we observed that limiting the number of NRFU contact attempts, as the Bureau has done in recent tests, is correlated with a high rate of unresolved cases—over 27 percent across the 2014, 2015, and 2016 tests. Additionally, in the 2016 test, cases were supposed to be limited to six attempts; even still, 10 percent of housing units received more than six attempts. Finally, the Bureau has not yet developed close-out procedures for these unresolved cases—another element that is not included in the cost estimate. Although the 2016 test sites are not nationally representative, we would be remiss if we did not report the potential magnitude of the as-yet unaccounted for costs in the 2020 Census life-cycle cost estimate.

We look forward to reviewing the Bureau's corrective action plan.

Appendix A: Objectives, Scope, and Methodology

This report addresses the Bureau's 2020 decennial program preparation and planning efforts. We initiated our evaluation of the 2016 Census Test to review (1) the effectiveness of the new management structure and (2) the effectiveness of the OCS to support NRFU operations. Our first objective was to determine whether the Bureau documented its decision to utilize the enumerator-to-supervisor ratios selected for the 2016 test. Our second objective was to determine whether the Bureau designed the test to (a) assess the effectiveness of the OCS in supporting supervisors during NRFU operations by comparing 2016 test results to results of previous tests and (b) determine the feasibility and effectiveness of a higher enumerator-to-supervisor ratio compared to the 2010 Census.

To accomplish our objectives we:

- Visited Houston and Los Angeles to observe field operations and interview Bureau officials;
- Interviewed Bureau officials in Suitland, Maryland;
- Reviewed training documentation;
- Identified and documented how the Bureau decided on the enumerator-to-supervisor ratios used at each location; and
- Analyzed computer-processed data to determine whether
 - enumerators were working during the times when respondents were most likely to be home,
 - alerts functioned as programmed,
 - supervisors responded to alerts,
 - enumerator visits exceeded the maximum number of attempts allowed on cases,
 - enumerators conducted proxy attempts, and
 - enumerators received paper responses.

To complete these tests, we reviewed data about alerts, enumerator contact attempts, enumerator assignments, and enumerator mileage and payroll. In conducting this analysis, we assessed the reliability of each dataset by performing reasonableness tests—for example: looking for missing data, calculation errors, data outside of valid time frames, data out of designated ranges, negative values in positive-only fields, and duplicate records. We also compared payroll data with enumeration data to determine whether there was a discrepancy between enumerator work days and enumerator pay days. We did not find any issues and consider the data sufficiently reliable for use in our analysis.

The following guidance and test plans were reviewed:

- *2016 Census Test Study Plan*;
- *2016 Census Test NRFU Supervisor Classroom Training Instructor Manual*; and
- *2016 Census Test NRFU Enumerator Classroom Training Instructor Manual*.

Further, we gained an understanding of the internal control processes significant to the evaluation objectives by interviewing officials at the Bureau and reviewing documentation and data for evidence of internal control procedures. Based on our review, we identified internal control weaknesses associated with the design of the 2016 Census Test. In addition, we noted several other internal control weaknesses related to the OCS and test processes that are included in this report.

This evaluation was conducted under the authority of the Inspector General Act of 1978, as amended, 5 U.S.C. App., and Department Organization Order 10-13, dated April 26, 2013. We conducted our fieldwork from May 2016 to November 2016 in accordance with the Quality Standards for Inspection and Evaluation (January 2012) issued by the Council of the Inspectors General on Integrity and Efficiency.

Appendix B: Agency Response



UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Office of the Director
Washington, DC 20233-0001

February 24, 2017

MEMORANDUM FOR: Carol Rice
Assistant Inspector General for Economic
And Statistical Program Assessment

From: John H. Thompson 
Director, U.S. Census Bureau

Subject: *2020 Census: 2016 Census Test Indicates the Current Life-Cycle
Cost Estimate is Incomplete and Underestimates Nonresponse
Followup Costs*
Draft Report

The attached comments are in response to your January 27, 2017, draft report titled “*2020 Census: 2016 Census Test Indicates the Current Life-Cycle Cost Estimate is Incomplete and Underestimates Nonresponse Followup Costs.*” The Census Bureau appreciates the opportunity to review and provide comments on this draft report.

Attachment

cc: Pam Moulder, ESA
Raul Cisneros, ESA



census.gov

Census Bureau Comments on OIG Draft Report:

"2020 Census: 2016 Census Test Indicates the Current Life-Cycle Cost Estimate is Incomplete and Underestimates Nonresponse Followup Costs"

February 2017

The U.S. Census Bureau appreciates the opportunity to comment on this draft report. We have no fundamental disagreements with any of the recommendations and will soon prepare a formal action plan to document the steps we will take in response to those recommendations.

While we concur with OIG's recommendation concerning the need to ensure accuracy and good documentation of our life-cycle cost estimates, we disagree with the detailed findings beginning on page 4, and with the title of the report derived from those findings, because we believe they are based on incorrect assumptions and extrapolations from our testing.

The statement that our life-cycle cost estimates fail to account for more than 10 million Nonresponse Followup (NRFU) cases – more than a 20 percent increase in our current estimated NRFU field cases of 46 million – cannot be supported. While this estimate is based on results in recent field tests, it does not factor in how those test sites and results are not comparable to the 2020 Census workload or environment.

Our life-cycle cost estimates are refined each year based on what is learned in our testing and in other research, but must always be modeled for the 2020 Census and its design and environment, not the design and environment of any one component test. The life-cycle cost model already includes the rates of non-interview after six NRFU contact attempts that we assume for a full and accurate count in the 2020 Census. It also includes funding for various situations that might require more field visits than expected. For example, a higher than expected number of households needing re-contact to resolve content or count discrepancies, or a higher number than expected of non-interviews at the conclusion of NRFU.

More specifically, section I of the draft report includes an extrapolation of around 30,000 NRFU visits beyond the sixth attempt during the 2016 Census Test to support its statement that "more than ten million potential NRFU contact attempts are unaccounted for in the life-cycle cost estimate." For the following reasons, the Census Bureau disputes the use of this extrapolation and does not concur that costs of this magnitude are missing from the life-cycle cost estimates.

1. The sample in the 2016 Census Test and other referenced field tests like the 2015 Census Test are not representative of the entire country. Recent 2020 Census field tests included parts of Houston, Los Angeles, and Phoenix, all containing significantly more hard-to-count areas than the Nation at large. Even if this test had an end-to-end focus rather than the much narrower focus it did have, results could not be extrapolated nationwide. In the 2020 Census, our expectation is that we will focus on working NRFU cases until completion. We expect, and have accounted for in our cost estimates, that in the 2020 Census some areas of the

country will take more than six attempts to resolve, while many others will take fewer than six attempts. The life-cycle cost model assumptions are nationwide rather than geographically dependent.

2. The field procedures used during the 2016 Census Test were not the final procedures that will be used in the 2020 Census. Unlike the 2018 End-to-End Census Test and ultimately the 2020 Census, the 2016 Census Test did not focus on closing out cases but rather focused on testing field management, including the enumerator-to-supervisor ratio. Other conditions were also not yet final:
 - Field staff were not using the final operational control system and other technology that will be utilized in the 2020 Census with optimized routing and case management capabilities.
 - Training must still be optimized for the final field procedures, which will reduce many of the repeated visits experienced in this test. A portion of these repeated contact attempts were reopening and reclosing the mobile case on the device rather than new knocks on the same doors, which will be reduced once operations and training are finalized.
 - We did not plan procedures to monitor and retain the highest performing enumerators to divert them on close-out of the most difficult cases. In small tests of this procedure, we were able to close out difficult cases expeditiously.
3. Taken together, these reasons highlight why extrapolating procedures in a test environment in 2016 is not analogous to refined and finalized field procedures we will have in the 2020 Census.
4. The response environment in 2016 is not equivalent to that of 2020. The draft report acknowledges on page 5 that “strategies intended to increase response rate” including “nationwide publicity, the Census partnership program, paid advertising” are not a part of field tests for the 2020 Census. This impacts not only self-response rate but also the willingness of nonresponding households to answer the door to an enumerator and respond to the Census test. For this reason, response rates and non-interview rates between mid-decade tests and the decennial census cannot be directly compared without more advanced data analysis.

The Census Bureau has included all housing units that are projected to be in this country in the 2020 Census in its life-cycle cost estimations, and there is no evidence that nearly all of these cases cannot be resolved with a six NRFU contact attempts. For the small fraction of cases that cannot be resolved in the field in the actual environment of the 2020 Census complete with finalized routing and assignment technology, we will utilize proxy responses and high-quality administrative records, and finally data imputation. All of this has already been accounted for in the life-cycle cost estimates.

Concerning the individual recommendations:

OIG Recommendation 1: Ensure that the 2020 Census life-cycle cost estimate accurately reflects all relevant cost factors and excluded costs are documented.

Census Bureau Response: The Census Bureau concurs with this recommendation. As subsequent iterations of the 2020 Census life-cycle cost estimates are prepared, we will continue our evolution of the cost estimates to reflect relevant cost factors and enhance our documentation of the life-cycle cost estimates.

OIG Recommendation 2: Designate appropriate personnel to independently verify that tests are properly designed to answer research questions.

Census Bureau Response: The Census Bureau accepts this recommendation. The Census Bureau has established the Decennial Research Objectives and Methods (DROM) Working Group. The DROM Working Group membership has a strong presence from the Census Bureau's Research and Methodology (R&M) Directorate including, but not limited to: the Associate Director for R&M, the Assistant Director for R&M, R&M's Senior Economist, and R&M's Senior Mathematical Statistician for Analytics. Other DROM members include R&M Center Chiefs, additional Senior Advisers, and representatives of other key stakeholder divisions. The DROM is accountable and responsible for deriving relevant decennial research objectives, advising on appropriate sound methods, and for producing meaningful and accurate reports for the 2020 Census.

OIG Recommendation 3: Work with the Office of General Counsel to develop enumerator scheduling guidance; and implement the most efficient NRFU enumeration scheduling practices.

Census Bureau Response: The Census Bureau agrees with this recommendation. This Census Bureau will take the necessary steps to investigate options for scheduling enumerator work to arrive at an implementable solution for the NRFU operation.

OIG Recommendation 4: Implement internal controls in the new operation control system that (a) prevent supervisors from ignoring alerts and inform managers that alerts were not responded to in a timely manner; and (b) provide supervisors with sufficient detail to resolve alerts.

Census Bureau Response: The Census Bureau agrees with this recommendation. As the Census Bureau has developed requirements for the new operational control system, we have considered approaches, such as no longer allowing alerts to expire and reports for managers that communicate data on the number of alerts not resolved in a timely manner. We will continue to modify the solution considering the OIG observations.

OIG Recommendation 5: Revise training to ensure field staff are adequately prepared to conduct proxy interviews and securely transmit paper questionnaires for processing.

Census Bureau Response: The Census Bureau agrees with this recommendation. As operational procedures evolve, whether based on 2016 Census Test experiences or as a result of further definition of the operational design, training and procedures will also evolve. An enumerator's job is multifaceted, necessitating training on many topics. Balancing training content against the critical components of an enumerator's job – while considering cost and schedule – will be key to our success.

011200072255