

Accessible Version

Highlights of [GAO-21-206R](#).

The Census Bureau has completed data collection operations for the 2020 Decennial Census. As the Bureau begins processing responses to deliver data for apportionment and redistricting, it will need to ensure the quality, accuracy, and protection of the data collected.

In recent years, GAO has identified challenges to the Bureau's ability to conduct a cost-effective count of the nation, including new innovations, acquisition and development of IT systems, and other challenges. In 2017, these challenges led GAO to place the 2020 Census on its High-Risk List.

Since 2007, GAO has made 113 recommendations specific to the 2020 Census. As of December 2020, 20 of the recommendations had not been fully implemented.

GAO was asked to provide regular updates on the 2020 Census. This report examines the cost and progress of key 2020 Census operations critical to a cost-effective enumeration, and early warnings that may require Census Bureau or congressional attention.

The Bureau provided technical comments that were incorporated as appropriate.

View [GAO-21-206R](#). For more information, contact J. Christopher Mihm at (202) 512-6806 or by email at mihmj@gao.gov and Nick Marinos at (202) 512-9342 or by email at marinosn@gao.gov.

2020 Census

The Bureau Concluded Field Work but Uncertainty about Data Quality, Accuracy, and Protection Remains

This report is the fifth in a series of updates on the Census Bureau's (Bureau) 2020 Census activities and operations. This update includes information from GAO's ongoing work on the conclusion of 2020 Census field operations, selected potential indicators of quality, and changes made to response processing operations as the Bureau produces its data products.

The Bureau changed the dates for completing the nonresponse follow-up (NRFU) operation and delivering data for apportionment several times between August 2020 and October 2020 in response to litigation. After receiving a ruling from the U.S. Supreme Court on October 13 that allowed it to stop data collection, the Bureau announced it would conclude NRFU on October 15 and deliver data for apportionment on or shortly after December 31.

At the conclusion of its data collection operations on October 15, the Bureau reported it had achieved a national enumeration rate of 99.98 percent of housing units. The Bureau reported 67.0 percent of enumerations came from self-response via internet, paper, or phone, and an additional 32.9 percent of households were enumerated during NRFU.

Data Collection Operations in the Field Have Ended, but Data Quality, Accuracy, and Protection Remain Uncertain

Alternative Data Collection Methods

When the Bureau cannot obtain census information directly from household members, either through self-response or a completed NRFU interview, it relies on alternative methods. The Bureau's reliance on these methods may provide insight into the quality of data collected:

- **Proxy responses.** The Bureau used proxy responses—information from a neighbor or other knowledgeable person, such as a landlord or building manager, about a household—to collect data on 24.1 percent (approximately 7.4 million, based on preliminary results) of occupied households in the NRFU workload, compared to 23.8 percent (approximately 6.8 million) in 2010. Proxy responses are generally lower quality than responses directly from a household.
- **Partial responses.** The Bureau may receive a partial response for a household through self-response or a NRFU interview. For some cases, enumerators in the field are directed to obtain, at a minimum, the status of whether the household is occupied, vacant, or not a household, and the number of people in the housing unit. The number of responses with this minimal amount of data can be an indicator of the quality of data collected. The Bureau has not yet calculated the number of partial responses it received, but plans to report on it in future operational assessments.
- **Administrative records.** The 2020 Census incorporated increased use of administrative records into its design, a major cost saving innovation. Use of these records leverages information people have already provided to the federal or state government, such as the Internal Revenue Service or prior census data. The Bureau used administrative records to resolve approximately 14 percent of households (about 8.4 million) in the NRFU workload, which was less than planned. However, the Bureau decided to use

administrative records that lack corroboration by a second source after NRFU began, introducing a data quality risk.

- **Imputation.** This statistical method draws on data from other household members, nearby households, and data on that household from past censuses and administrative records. The Bureau uses imputation to create records for housing units that appear occupied, but for which no other information is available. It has been used in some form since the 1940 Census.

Difficulty Completing NRFU in Some Local Areas

The census is a local endeavor and the Bureau experienced challenges completing NRFU in some local areas including difficulty hiring enumerators and accessing rural and tribal areas. Other challenges included high rates of COVID-19 and natural disasters.

To address these challenges, the Bureau instituted financial awards for enumerators who maximized hours and completed a set number of cases per hour. The Bureau also enumerated some areas by phone and used travel teams of enumerators, offering financial awards for those willing to travel to certain areas.

When the Bureau left the field on October 15, 10 of the Bureau's 248 area census offices fell short of completing 99 percent of their NRFU workload, one of the Bureau's stated indicators of completion.

Less Time to Ensure Accuracy during Response Processing Operations

To deliver data for apportionment on December 31, the Bureau will have only 77 days to complete response processing, an operation that was designed to take 153 days. To complete response processing in fewer days, the Bureau made changes to its process, including locking down its Master Address File prior to the end of data collection and shortening the amount of time for reviews by subject matter experts in the Bureau's statistics divisions.

The Bureau is also prioritizing tasks needed to produce apportionment counts rather than simultaneously preparing redistricting data, which involves more data elements. In doing this, the Bureau will create two separate analyses for the separate output files, which differs from its plan to produce a single analysis to support both output files.

In compressing its response processing, the Bureau also faces increased risk that system defects or other information technology issues may go undetected, affecting the quality and accuracy of the count. Additionally, the Bureau will have less time to address issues that arise.

Work Remains to Protect Data Privacy

The Bureau reported progress in implementing a disclosure avoidance technique to protect the confidentiality of its respondents' data in its publicly-released statistical products. However, the Bureau still has work remaining before it implements disclosure avoidance on its data products, and final decisions regarding the implementation have yet to be made.



441 G St. N.W.
Washington, DC 20548

December 9, 2020

Congressional Requesters

2020 Census: The Bureau Concluded Field Work but Uncertainty about Data Quality, Accuracy, and Protection Remains

In response to the Coronavirus Disease 2019 (COVID-19) national emergency, the Census Bureau (Bureau) adjusted its plans for the 2020 Census several times.¹ In addition, the Bureau faced uncertainty regarding executive and congressional action and the outcome of ongoing litigation over the past 5 months. The Bureau ultimately finished data collection on October 15, 2020, and stated at that time it was endeavoring to deliver apportionment data by or as close to the statutory deadline of December 31, 2020, as possible.

You asked us to provide regular updates on the implementation of the 2020 Census. For these updates, we review the cost and progress of key 2020 Census operations critical to a cost-effective enumeration and early warnings that may require Bureau or congressional attention. For this correspondence—the fifth in a series of products—we focused on census operations since our last report in August 2020, including the completion of field operations, selected potential indicators of quality, and changes made to response processing operations.²

To describe the status of key operations for the 2020 Census and describe major trends and early warning signs, we reviewed Bureau-provided data on cost and progress of key operations and compared those data with the Bureau's plans and Bureau-determined target dates and metrics. We determined those data were sufficiently reliable for the purposes of our reporting objectives by interviewing Bureau staff about the information technology (IT) systems used. We interviewed Bureau officials to gather additional information on the status and progress of these key operations.

In addition, to obtain a ground-level perspective on the conduct of key field activities, we surveyed the Bureau's entire population of 248 area census office (ACO) managers 6 times during the 2020 Census, including in late February to early March, early April, late May, late June to early July, late August, and early October. The response rates were 71, 75, 76, 72, 67, and 69 percent, respectively. We also reviewed open-ended responses provided by the ACO managers as part of this survey.

We also included information from our ongoing work related to the Bureau's IT system implementation and data protection activities for the 2020 Census. We collected and reviewed documentation on the status of disclosure avoidance activities, such as milestone schedules and differential privacy updates. We also interviewed relevant agency officials about their plans to implement disclosure avoidance methods for 2020 Census data products.

¹For more information see GAO, *2020 Census: Census Bureau Needs to Assess Data Quality Concerns Stemming from Recent Design Changes*, [GAO-21-142](#) (Washington, D.C.: Dec. 3, 2020).

²For the prior reports in this series, see GAO, *2020 Census: Recent Decision to Compress Census Timeframes Poses Additional Risks to an Accurate Count*, [GAO-20-671R](#) (Washington, D.C.: Aug. 27, 2020); *2020 Census: COVID-19 Presents Delays and Risks to Census Count*, [GAO-20-551R](#) (Washington, D.C.: June 9, 2020); *2020 Census: Initial Enumeration Underway but Readiness for Upcoming Operations is Mixed*, [GAO-20-368R](#) (Washington, D.C.: Feb. 12, 2020); and *2020 Census: Status Update on Early Operations*, [GAO-20-111R](#) (Washington, D.C.: Oct. 31, 2019).

We conducted this performance audit from September 2020 to December 2020 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Background

Past Work on the 2020 Census

In 2017, we designated the 2020 Census as a high-risk area and added it to our High-Risk List.³ The 2020 Census remains on the list, as new innovations, acquisition and development of IT systems for the 2020 Census, and other challenges we have identified in recent years—such as the development of risk mitigation plans—raise serious concerns about the Bureau’s ability to conduct a cost-effective enumeration.⁴ Since 2007, we have made 113 recommendations specific to the 2020 Census to help address these risks and other concerns.

The Department of Commerce (Commerce) has generally agreed with these recommendations and has taken action and made progress to address them. However, as of December 2020, 20 of the recommendations have not been fully implemented and 10 of these are designated as priority recommendations.⁵

The Bureau’s Timelines Have Changed Multiple Times

In August 2020 we reported that, after pausing its operations in response to the COVID-19 pandemic, the Bureau began to resume selected operations in May 2020.⁶ In April 2020, Commerce requested that Congress extend its statutory reporting dates by 120 days. The extensions would have allowed the Bureau to deliver data for apportionment to the President by April 30, 2021.⁷

However, on August 3, the Bureau announced that it would deliver the apportionment counts by the statutory deadline of December 31, 2020, and that it would accelerate completion of its data collection and data processing operations to do so. Since that time, the dates for the Bureau to complete nonresponse follow-up (NRFU) and deliver apportionment data have changed several times in response to litigation regarding these accelerated time frames (see fig. 1).

³GAO’s high-risk program identifies government operations with vulnerabilities to fraud, waste, abuse, and mismanagement, or in need of transformation to address economy, efficiency, or effectiveness challenges.

⁴GAO, *High-Risk Series: Substantial Efforts Needed to Achieve Greater Progress on High-Risk Areas*. [GAO-19-157SP](#) (Washington, D.C.: March 2019) and *High-Risk Series: Progress on Many High-Risk Areas, While Substantial Efforts Needed on Others*. [GAO-17-317](#) (Washington, D.C.: Feb. 15, 2017).

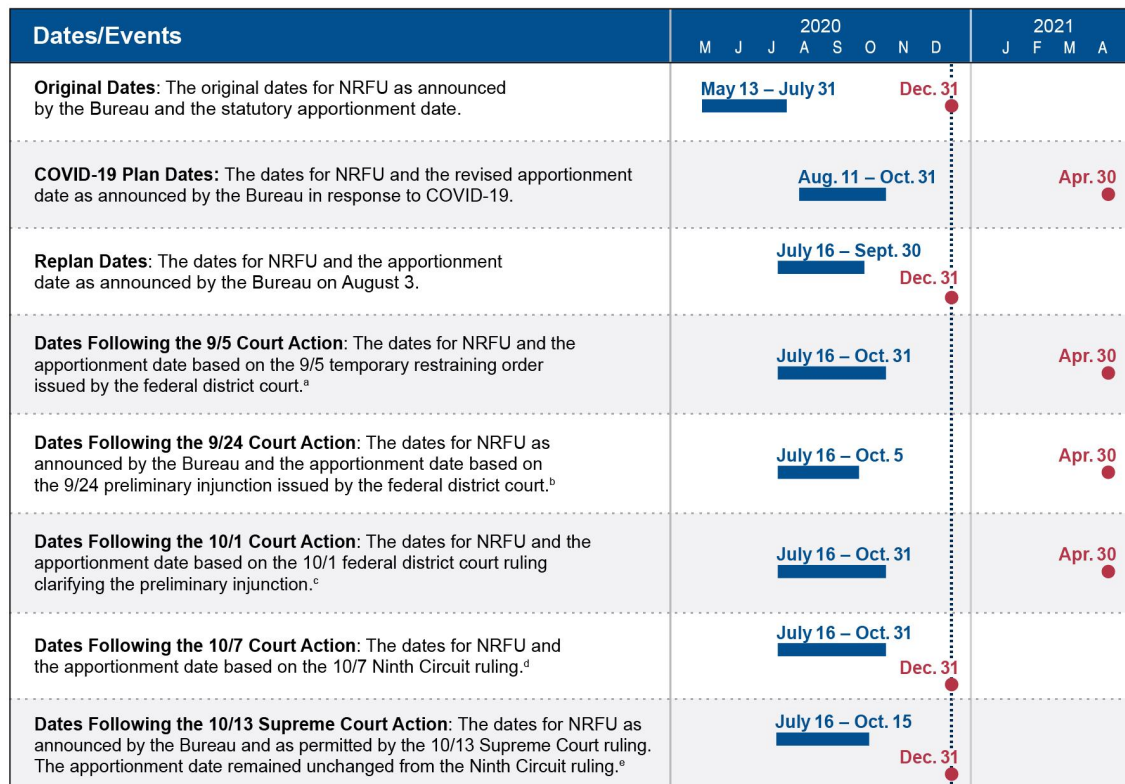
⁵Priority recommendations are those that GAO believes warrant priority attention from heads of key departments or agencies. They are highlighted because, upon implementation, they may significantly improve government operations; for example, by realizing large dollar savings; eliminating mismanagement, fraud, and abuse; or making progress toward addressing a high-risk or fragmentation, overlap, or duplication issue.

⁶[GAO-20-671R](#).

⁷While legislation providing this relief has been introduced in Congress, the legislation has not been enacted into law. One bill, The Heroes Act, has passed the House of Representatives. The Heroes Act, H.R. 6800, 116th Cong. (2020). See also Fair and Accurate Census Act, H.R. 7034, S. 4048, 116th Cong. (2020) and 2020 Census Deadline Extensions Act, H.R. 8250, S. 4751, 116th Cong. (2020).

On October 13, the U.S. Supreme Court issued a ruling that permitted the Bureau to end data collection. The Bureau announced it would end data collection operations 2 days later on October 15, 2020. For more information on this litigation, see enclosure I.

Figure 1: Dates for Nonresponse Follow-up (NRFU) and Delivery of Apportionment Data



- Date for Delivering Apportionment Data
- Statutory Deadline for Delivering Data for Apportionment

Source: GAO analysis of Census Bureau and court documents. | GAO-21-206R

Data for Figure 1: Dates for Nonresponse Follow-up (NRFU) and Delivery of Apportionment Data

| Dates/Events | Information for the NRFU Dates Bars | Working Date for Delivering Apportionment Date |
|---|---|--|
| Statutory Deadline for Delivering Data for Apportionment 12/31/20 | -- | -- |
| Original Dates: The original dates for NRFU as announced by the Bureau and the statutory apportionment date. | Start Date: 5/13/20 End Date: 7/31/20 | 12/31/20 |
| COVID-19 Plan Dates: The dates for NRFU and the revised apportionment date as announced by the Bureau in response to COVID-19. | Start Date: 8/11/20 End Date: 10/31/20 | 4/30/21 |
| Replan Dates: The dates for NRFU and the apportionment date as announced by the Bureau on August 3. | Start Date: 7/16/20 End Date: 9/30/20 | 12/31/20 |

| Dates/Events | Information for the NRFU Dates Bars | Working Date for Delivering Apportionment Date |
|--|--|---|
| Dates Following the 9/5 Court Action: The dates for NRFU and the apportionment date based on the 9/5 temporary restraining order issued by the federal district court. ^a | Start Date: 7/16/20 End Date: 10/31/20 | 4/30/21 |
| Dates Following the 9/24 Court Action: The dates for NRFU as announced by the Bureau and the apportionment date based on the 9/24 preliminary injunction issued by the federal district court. ^b | Start Date 7/16/20 End Date: 10/5/20 | 4/30/21 |
| Dates Following the 10/1 Court Action: The dates for NRFU and the apportionment date based on the 10/1 federal district court ruling clarifying the preliminary injunction. ^c | Start Date: 7/16/20 End Date: 10/31/20 | 4/30/21 |
| Dates Following the 10/7 Court Action: The dates for NRFU and the apportionment date based on the 10/7 Ninth Circuit ruling. ^d | Start Date: 7/16/20 End Date: 10/31/20 | 12/31/20 |
| Dates Following the 10/13 Supreme Court Action: The dates for NRFU as announced by the Bureau and as permitted by the 10/13 Supreme Court ruling. The apportionment date remained unchanged from the Ninth Circuit ruling. ^e | Start Date: 7/16/20 End Date: 10/15/20 | 12/31/20 |

^aOn September 5, a federal district court issued a temporary restraining order that prevented the Bureau from using the Replan Dates announced August 3.

^bOn September 28, the Bureau announced that NRFU would end on October 5 despite the September 24 preliminary injunction enjoining the Bureau from operating under the August 3 Replan. The Bureau said it made this decision in order to be able to deliver apportionment data by December 31, in case the court's decision was reversed.

^cIn response to the Bureau's plan to end NRFU on October 5, the federal district court issued a clarification of its preliminary injunction stating that the Bureau could not end NRFU before October 31. The Bureau announced it would use the court-imposed date on October 2.

^dOn October 7, the Ninth Circuit Court of Appeals issued a ruling agreeing with the district court that the Bureau could not end NRFU before October 31. However, the circuit court reversed the district court with respect to the December 31 apportionment date, finding that the Bureau could attempt to deliver the apportionment numbers by the end of the year.

^eOn October 13, the Bureau announced that NRFU would end on October 15 as a result of the October 13 Supreme Court ruling that the Bureau was not required to continue operations until October 31. Additionally, the Bureau announced it planned to deliver data for apportionment by or as close to the statutory date of December 31 as possible.

The Bureau Completed Data Collection for the 2020 Census

The Bureau ended its data collection—including self-response and field data collection—on October 15.⁸ At the close of these operations, the Bureau reported that it achieved a national enumeration rate of 99.98 percent.⁹ According to the Bureau, 67.0 percent (99.0 million) of housing units self-responded via internet, phone, or paper by that date. The Bureau reported that during NRFU—the largest data collection operation in the field—the Bureau enumerated an additional 32.9 percent of housing units (48.6 million, based on preliminary data).¹⁰

In July, for the NRFU operation, the Bureau started sending enumerators to attempt to enumerate housing units that had not yet responded to the 2020 Census. By October 15, the Bureau reported that it had completed 99.93 percent of its NRFU workload (almost 64.1 million households), which includes cases of occupied, vacant, and nonexistent housing units as well as cases where the Bureau returned to households with completed interviews to verify information through an additional interview.¹¹

Bureau officials told us that they took a number of steps to ensure that the national NRFU workload completion goal was reached by October 15. In August, we reported that the Bureau introduced financial awards nationally to reward enumerators who maximized productivity by completing a set number of cases per hour and working increased hours.¹² As of October 15, the Bureau reported \$70.3 million in expenses for enumerator awards. Bureau officials also said the Bureau regularly communicated about the end of NRFU with all field and office staff.

According to the Bureau, it experienced higher enumerator productivity than it expected, which aided it in reaching its workload goal.¹³ According to Bureau reporting, as of October 15, a cumulative 1.92 cases were completed per hour, exceeding the Bureau's planned 1.55 cases completed per hour, which the Bureau attributed in part to new technology and households self-responding after receiving a notice of visit from a NRFU enumerator.

However, our early October survey of ACO managers found that fewer than a quarter of ACO manager respondents reported satisfaction with the accuracy and efficiency of the Bureau's IT system that optimized both the assignment and routing of cases for enumerators (22 percent). In early October, ACO managers' comments during this time expressed concerns regarding

⁸Self-response was available across the nation through October 15, 2020, until 11:59 p.m. Hawaii Standard Time, which was 5:59 a.m. Eastern Daylight Time on October 16, 2020.

⁹To capture data from data collection operations ending later than October 15 Eastern Daylight Time due to different time zones, the figures reported here and in subsequent statements are from data collection results reported by the Bureau on October 19, unless noted otherwise.

¹⁰Reported percentages pertaining to 2020 Census response rates are based on preliminary and working data. According to Bureau officials, the Bureau will provide final data after reconciling all operational data to create the quality metrics that will accompany apportionment data.

¹¹The Bureau considers a NRFU case completed when (1) a household self-responds after being added to the NRFU workload; (2) the Bureau, after visiting the household at least once, has high-quality administrative records such as Internal Revenue Service or Social Security records that it can use to enumerate the household; and (3) the Bureau completes an acceptable field interview with the household. The Bureau also considers a NRFU case completed if two enumerators determine that an address is vacant or nonexistent by observation or talking to a proxy.

¹²[GAO-20-671R](#).

¹³Throughout the operation, the Bureau reported preliminary numbers of cumulative cases completed per hour. On October 26, Bureau officials stated that actual productivity data may take up to months to reconcile.

enumerators reporting performance issues, system slowness, and inefficiencies in the optimization of assigning and routing of enumerator work.¹⁴

Bureau officials stated that they believed the IT system worked well based on the higher-than-expected number of cases completed per hour, as discussed earlier. However, they also noted that dissatisfaction with the system may have been due to, among other things, ACO managers and enumerators not fully understanding how the system worked, which led to expectations that were not met.

The Bureau also completed other field data collection operations such as Group Quarters (GQ) and Service Based Enumeration (SBE) in a challenging environment involving COVID-19, wildfires, hurricanes, and civil unrest activities. The Bureau reported it concluded its GQ operation on September 3.¹⁵ According to the Bureau, this operation collected information from 271,983 GQ facilities including nursing homes, prisons, and on-campus college housing. The Bureau also reported it completed its SBE operation by enumerating in over 53,000 locations, including its count of individuals living in almost 37,000 Targeted Non-Sheltered Outdoor Locations (TNSOL) on September 30.¹⁶ These operations target persons experiencing homelessness and count individuals at various locations including shelters, soup kitchens, and mobile food vans.

Data Collection Operations in the Field Have Ended, but Data Quality, Accuracy, and Protection Remain Uncertain

Bureau Reliance on Alternative Data Collection Methods May Be a Key Indicator of Data Quality

According to the Bureau, questionnaires completed by household members provide a higher level of quality of census data whether they are submitted through self-response or completed with a census enumerator. Nationwide, preliminary results indicate that approximately 79 percent of households (116.1 million households) submitted their census data by self-responding via the internet, phone, or paper form, or through a completed NRFU interview with a household member. In 2010, 84 percent of households (approximately 108.8 million) self-responded or completed a NRFU interview with a household member.

In the remaining 21 percent (31.04 million households, based on preliminary data) where the Bureau did not obtain a completed questionnaire from a household member, it relied on alternative methods to create a complete record, either in the field by interviewing a proxy, during or after field operations are complete through the use of administrative records, or during response processing using statistical imputation.

In October 2020, the 2020 Census Quality Indicators Task Force at the American Statistical Association released a report on quality indicators and analyses that could be produced by the Bureau to assess the performance of the 2020 Census operations, including analysis across

¹⁴In late August, fewer than half of ACO manager respondents reported satisfaction with the accuracy and efficiency of the Bureau's IT system for assigning cases to enumerators (40 percent) as well. In late August, comments from ACO managers reflected similar concerns.

¹⁵The Bureau continued Late GQE through October 15. Late GQ enumeration allows for stakeholder identification and enumeration of GQs that may have been missed during the main GQ timeframe.

¹⁶According to Bureau officials, the Bureau planned to end TNSOL on September 24 and finished the operation in all planned locations except for San Benito, California which was enumerated as additional workload on September 30.

geographic regions and comparison with 2010 data.¹⁷ The report made several recommendations, including that these indicators be made readily available to the public to ensure understanding of the quality of the census. In November, the Bureau's Census Scientific Advisory Committee also recommended that these indicators be made available.¹⁸

Looking at the rates at which the Bureau uses each of these alternative methods of data collection may provide some insight into the quality of the data collected as part of the 2020 Census.¹⁹ Nationwide rates provide a high level indication of overall census quality. However, and very importantly, the Bureau also would need to examine the rates at which the Bureau uses these alternative methods at smaller levels of geography and by demographic group to obtain a complete picture.²⁰

- **Proxy responses.** The Bureau uses proxy responses—information from a neighbor or other knowledgeable person, like a landlord or building manager, about a household—to collect data when a resident of the household is not available or cannot be found during the NRFU operation. The Bureau prompts enumerators to find a proxy after the third unsuccessful contact attempt to a housing unit considered occupied.

Proxy responses are also allowable after one unsuccessful contact attempt to an address considered vacant or nonexistent. For the 2020 Census, preliminary results indicate that the Bureau enumerated approximately 7.4 million occupied households using proxies, or 24.1 percent of occupied households in the NRFU workload. By comparison, for the 2010 Census, the Bureau reported that a proxy was the respondent for approximately 23.8 percent of occupied households in the NRFU workload (6.8 million households).

Further, according to preliminary results, the Bureau resolved an additional 13.3 million NRFU cases in the 2020 Census as vacant or nonexistent using proxies. According to the Bureau, these cases are almost always resolved via proxy since there is not a household to respond.

- **Partial responses.** The Bureau receives incomplete questionnaires that are submitted through self-response and incomplete NRFU interviews in cases where an enumerator is unable to complete the census questionnaire for a household with either a household

¹⁷American Statistical Association 2020 Census Quality Indicators Task Force, *2020 Census Quality Indicators*, (Alexandria, VA.: October 2020).

¹⁸The Census Scientific Advisory Committee is an advisory body to the Director of the Bureau and composes recommendations on major programs, such as the decennial census. The members advise the Bureau on the uses of scientific developments in, among other things, statistical data collection and statistical analysis, as they pertain to the full range of Census Bureau programs and activities.

¹⁹In addition to the information presented here on the use of administrative records and proxies in closing NRFU cases, the Bureau will likely provide additional information on the use of these methods in the Operational Assessments produced after the 2020 Census. For example, in the 2010 Census Operational Assessments, the Bureau provided information on the number of housing units enumerated by proxies, number of housing units with incomplete responses, and the imputation rates for each household and person characteristic. The Bureau will begin producing these assessments in summer 2021.

²⁰According to Bureau officials, it has not provided data on use of these indicators at smaller levels of geography or by demographic group in previous censuses. Rather, since 1980, the Bureau has used its Post-Enumeration Survey to produce measures of coverage errors for geographic areas and demographic groups. Bureau officials told us they understand the importance of providing this information below the national level to the public in as near real time as possible this decennial census and are working to provide that data.

member or a proxy.²¹ For some of these cases, while in the field during the final phase of NRFU, enumerators are directed to try to obtain, at a minimum, an indication of the household status (vacant, not a housing unit, or occupied) and the number of the people in the housing unit by talking to a proxy.

The total number of responses with this minimal amount of data can be an indicator of the quality of data collected. Additionally, data on personal characteristics, which are not collected in minimal response situations, serve an important purpose.²² The Bureau has not yet calculated the number of partial responses it received, but plans to report on it in future operational assessments, the first of which are expected in 2021.

After the Bureau has completed its field data collection efforts, it generally finds that, for a small proportion of responses, (1) some households are missing responses altogether, or (2) some household responses include answers that are incomplete or conflict with one another.²³ For the 2020 Census, the Bureau used administrative records and statistical imputation to resolve some of these missing responses.

- **Administrative records.** The Bureau incorporated increased use of administrative records into the design of the of the 2020 Census as a major cost saving innovation and to improve the overall quality of the data. Administrative record data are information from federal and state governments and third party vendors that people have already provided. For example, the Bureau planned to use information from the United States Postal Service, the Internal Revenue Service, the Social Security Administration, the 2010 Census and the American Community Survey to resolve households that received an unsuccessful NRFU contact attempt and were identified as vacant or not a housing unit. The Bureau also used administrative records to enumerate some occupied households if a NRFU contact attempt was not successful.

Based on preliminary data, the Bureau used administrative records to reduce the number of NRFU contact attempts for fewer households than expected. According to preliminary Bureau data, after a single NRFU attempt, the Bureau was able to use administrative records that met a certain quality threshold to resolve 13 percent of NRFU cases (about 8 million cases) that were either occupied, vacant, or not a housing unit. Based on the administrative records modeling the Bureau did in 2019, it could have potentially used administrative records to resolve about 22.5 percent of NRFU cases after one NRFU contact attempt.

In the final phases of the NRFU operation, cases in which households had received the maximum number of contact attempts (typically six) were reopened. For these households, if the Bureau was still unable to conduct an interview, then the Bureau used administrative

²¹The Bureau determines whether or not a partial response is considered sufficient based on a set of criteria. According to the Bureau, it needs to collect a number of pre-defined specific combinations of data elements during field interviews in order to consider the response complete. Bureau officials said this criteria is sensitive.

²²Federal agencies use the Bureau's data on race and ethnicity to help monitor compliance with anti-discrimination provisions, such as those included in the Voting Rights Act. Pub. L. No. 89-110, 79 Stat. 437 (1965). Section 2 of the act protects minority population voting rights, and the Department of Justice and the courts use the Bureau's data in determining violations of the act. 52 U.S.C. § 10301(a).

²³In 2010, there were 492,515 occupied housing units in the NRFU workload for which enumerators did not know the population count. The Bureau conducted a supplemental NRFU operation, NRFU Residual, to obtain more information from households with no population count and/or partial responses to the questionnaire. After conclusion of this operation, there were 17,869 housing units that still had an unknown population count and were not marked as a refusal.

records that met the Bureau's quality standards if available. According to Bureau planning documents, while these administrative records used at the end of NRFU do not meet the threshold for early removal after the first visit, they do provide adequate data to support apportionment. Additionally, the Bureau decided to expand this use of administrative records and close some cases on the last day of NRFU, regardless of the number of contact attempts. For these households, the Bureau was able to use administrative records to resolve an additional approximately 450,000 cases (0.73 percent of the NRFU workload, according to preliminary data) that in past censuses would have needed to be resolved with statistical imputation. In sum, according to preliminary data, the Bureau used administrative records to close NRFU cases for 8.4 million households (approximately 14 percent of the NRFU workload), or approximately 5.6 percent of households nationwide.

The Bureau's original plan had been to use an administrative record to determine the population of a household only when corroborated by a second administrative record. However, as part of its plans to meet the statutory date for delivering apportionment data on December 31, the Bureau decided, after NRFU had already begun, to enumerate some cases using sole sources of administrative records when available.

This decision introduced a data quality risk, since there is no corroborating source. Bureau officials said they will be examining the effects on quality of the late design changes to use of administrative records. They also said the risk to overall data quality should be mitigated by the fact that the Bureau only used sole source records for households that lack a population count during the final stage of NRFU. At that point, most households had already been counted and households lacking a population count would otherwise require count imputation. The Bureau is still working on an estimate of its use of sole source records.

- **Imputation.** In cases where no administrative records are available to complete questionnaires for households with missing data, the Bureau uses both count and characteristic imputation to complete records. Imputation is a statistical technique that draws on data from other household members, nearby households, and data on that household from past censuses and administrative records (see textbox). The Bureau also uses imputation to create records for housing units that appear occupied, but for which no other information is available. The Bureau has used some form of imputation since the 1940 Census, and it reported on the use of imputation in its operational assessment reports for the 2010 Census.

Three Types of Characteristic Imputation Used by the Census Bureau:

- **Assignment.** If a response for a data item is either missing or not consistent with other responses in the questionnaire, and an item value can be determined based on other information provided for the same person, the Bureau will input a value for that item. For example, if a person's race is missing, but a write-in response for the Hispanic origin field identifies race, then the race can be assigned.
- **Allocation.** If a response for a data item is either missing or not consistent with other responses in the questionnaire and an item value cannot be determined based on other information provided for that person, the Bureau will input a response from another person within the housing unit or from a person in a nearby housing unit.
- **Substitution.** If the characteristics for all the persons in a household are missing, the Bureau will use a nearby household with complete person data to represent the person-level items for the persons needing substitution.

The Bureau uses count imputation to resolve missing household responses following data collection. It draws data from similar nearby households to determine whether a housing unit exists, whether it is occupied, and, if so, by how many people.

Source: GAO analysis of Census Bureau documents | GAO-21-206R

The Bureau Had Difficulty Completing NRFU in a Few Areas

While the national response rate is important, the decennial census is a local endeavor. As operations progressed, the Bureau experienced challenges completing NRFU in some local areas. In late August, the Bureau began designating local areas that were not meeting their enumeration workload goals as challenge areas. According to the Bureau, areas designated as challenge areas were those that experienced difficulties with hiring, accessing rural and tribal areas, the effects of high rates of COVID-19, and natural disasters such as hurricanes and wild fires.

As of October 15, the last day of data collection, 10 of 248 ACOs had not completed 99 percent of their reported workload—one of the Bureau’s stated indicators of completion (see table 1).²⁴

Table 1: Area Census Offices that Completed Less Than 99 percent of Their Nonresponse Follow-up (NRFU) Workload

| Area census office | State | Percent of NRFU workload complete | Number of open cases |
|--------------------|---------------|-----------------------------------|----------------------|
| Shreveport | Louisiana | 93.54 | 22,588 |
| Manhattan 2 | New York | 98.49 | 4,192 |
| Window Rock | Arizona | 98.74 | 1,129 |
| San Francisco | California | 98.75 | 1,999 |
| Chicago South | Illinois | 98.86 | 1,929 |
| Quincy | Massachusetts | 98.88 | 1,104 |
| Sunnyvale | California | 98.89 | 1,803 |
| New Haven | Connecticut | 98.93 | 1,876 |
| Vista | California | 98.94 | 1,430 |
| Colorado North | Colorado | 98.96 | 2,823 |

Source: GAO analysis of Census Bureau reports. | GAO-21-206R

Note: The data above are as of October 19, 2020 which accounts for delays in reporting. However, the NRFU operation ended on October 15, 2020.

Challenges experienced by these offices varied. For example, the Bureau reported that the Shreveport ACO had been unable to access an area that was restricted because of the damage from Hurricane Delta. The Bureau also said it experienced a number of challenges in Arizona including closures of tribal lands due to COVID-19.²⁵ Bureau officials told us that data collection ceased at ACOs on October 15, regardless of whether they had completed 99 percent of their workload. Open cases were completed through imputation and administrative records.

Hiring in general was a challenge for the Bureau. In August, we reported that the Bureau planned to hire up to 435,000 enumerators to conduct NRFU nation-wide. The Bureau reported that as of October 15, it had hired 373,784 NRFU enumerators nation-wide. Bureau officials stated that while some enumerators quit because they were worried about catching COVID-19,

²⁴In calculating the percentage of NRFU workload complete, we did not count cases where enumerators had not obtained an interview or proxy data after six attempts as completed cases. In the Bureau’s public data, the Bureau counts these as completed cases, yielding two ACOs with less than 99 percent of their NRFU workload complete—Shreveport, Louisiana and Window Rock, Arizona.

²⁵We use the term tribal lands to include both reservation and off-reservation trust lands.

other enumerators worked more hours to make up for the hiring gap. They also said enumerators were more productive than expected.

To address local hiring challenges, Bureau officials noted that the Bureau increased the number of enumerators working in challenge areas through the use of travel teams, composed of experienced enumerators from other areas. The Bureau reported that it began sending travel teams to challenge areas in Oklahoma and Montana in September, and ultimately deployed travel teams in all U.S. states and territories except Hawaii and Puerto Rico.²⁶ The Bureau made pay awards available to enumerators willing to travel to states with especially low numbers of enumerators working, such as Alabama. A Bureau official told us that, throughout the operation, approximately 26,000 enumerators traveled from other areas and regions.

In addition to hiring, the Bureau also experienced difficulties with gaining entry to some tribal land areas at the start of NRFU because they were closed due to COVID-19.²⁷ A census field supervisor we interviewed told us that evening and weekend lockdowns due to COVID-19 made it difficult for enumerators to get people to open their doors. Even after tribes reopened their lands to enumerators, the Bureau had difficulties completing NRFU, especially in areas that included additional restrictions on completing work within their boundaries, such as requesting that only tribal members do the work or accompany census workers within their boundaries.

The Bureau worked with individual tribal nations and, in some cases, the Bureau obtained approval for non-American Indian enumerators to work on the tribal land either with or without escorts. In other cases, the Bureau moved American Indian enumerators to cover a different tribe's lands as a contingency to complete NRFU enumeration. The Bureau reported that as of October 16, the Bureau had completed 99.77 percent of the NRFU workload on American Indian and Alaska Native lands.

Natural disasters such as wildfires and hurricanes also caused difficulties. In areas that experienced natural disasters, the Bureau used telephone contact to reach displaced persons. Bureau officials told us that enumerators made over 10 million phone attempts and completed about 1.2 million cases by phone.²⁸ This enabled the Bureau to continue to employ enumerators that were displaced due to natural disasters and complete NRFU cases.

Responses to our early October survey of ACO managers found that fewer than half of ACO manager respondents reported satisfaction with the time and resources provided to meet production goals while maintaining data quality (40 percent). In early October, ACO manager comments during this time expressed concerns regarding completing cases in a compressed time frame; concerns with staff due to the quality, availability, or number of staff; and concerns with the accuracy of the completed cases.²⁹

²⁶According to the Bureau, Puerto Rico did not have any travelers from other regions.

²⁷For more information on how American Indian/Alaska Native people were affected by COVID-19 see GAO, *COVID-19: Opportunities to Improve Federal Response and Recovery Efforts*, [GAO-20-625](#) (Washington, D.C.: June 25, 2020).

²⁸Enumerators called nonresponding households and offered to complete their questionnaire for them over the phone. In cases where the enumerator did not reach the respondent, the enumerator left a voicemail to remind the respondent of the ways in which they could complete the census questionnaire. Telephone contact also includes calls made by enumerators when they located a phone number in the field, such as for a real estate agent or building manager.

²⁹In late August, fewer than half of ACO manager respondents reported satisfaction with the time and resources provided to meet production goals while maintaining data quality (40 percent) as well. In late August, comments from

The Bureau Has Less Time to Ensure Accuracy of Data during Response Processing Operations

As of October 16, the Bureau planned to shorten response processing operations to deliver data for apportionment on or shortly after December 31. These activities were initially designed to take 153 days. If the Bureau plans to deliver data for apportionment on December 31, it will have only 77 days to complete response processing.

During response processing, the Bureau improves the accuracy of the data collected; checks for and resolves duplicate, inconsistent, and incomplete responses; and uses administrative records to supplement response data. According to Bureau officials, some components of response processing must be performed sequentially, since each component builds on the previous component.

To complete response processing in fewer days, the Bureau had to make changes to its response processing operation. For example, Bureau officials told us that they locked down the Master Address File (MAF) of all living quarters by September 27, 2020, prior to the end of NRFU and the self-response period. Bureau officials said that any new addresses detected after that date would not become part of the MAF for apportionment or redistricting purposes. The Bureau estimated there were 134,461 such addresses.

Bureau officials also told us that they did not remove any levels of review by subject matter experts in the Bureau's statistics divisions of the accuracy of state-by-state counts, but they did shorten the amount of time for those reviews. We have asked the Bureau for information on how time frames have changed for response processing, including changes to the amount of time available for subject matter expert review.³⁰ The Bureau has not yet provided that information. In 2010, subject matter expert reviews resulted in at least one re-run of the analyses for all 50 states, Puerto Rico, and the District of Columbia prior to the release of the output file for apportionment.

The Bureau said it would prioritize tasks needed to produce apportionment counts rather than simultaneously preparing redistricting data, which involves more data elements. To accomplish this, the Bureau is conducting two separate analyses for the separate output files for apportionment and redistricting. This differs from the Bureau's plan to produce one analysis that would support both output files. We requested, but have not yet received, information on how the Bureau will ensure accuracy across these output files.

Before response processing began, the Bureau reported that it completed testing of the IT systems needed to deliver data for apportionment. However, the Bureau reduced the amount of time it took to conduct this testing. Specifically, the Bureau compressed its testing schedule in order to complete testing in October 2020 instead of January 2021, as previously planned. As a result, the Bureau faces an increased risk that system defects or other issues may go undetected and could affect the quality and accuracy of the count. In addition, the shortened time for response processing presents additional risk because there will be less time available to address system defects or other issues.

more than 20 ACO managers reflected similar concerns regarding inability to complete NRFU in time, insufficient staff in some locations, and resulting inaccuracy in hard-to-count areas.

³⁰The Bureau conducts its subject matter expert review with other divisions within the Census Bureau including the Social, Economic, and Housing Statistics Division, the Population Division, and the Decennial Statistical Studies Division.

The Bureau Has Work Remaining to Protect the Privacy of Respondent Data

To protect the privacy of respondent data as required by statute,³¹ the Bureau reported progress in implementing a disclosure avoidance technique, known as differential privacy,³² to its publicly-released statistical products to protect the confidentiality of its respondents and their data.

The Bureau continues to solicit feedback and make improvements to its disclosure avoidance approach, and its implementation of differential privacy, based on input it receives from data users and the public. In September 2020, the Bureau's Census Scientific Advisory Committee provided the Bureau with 11 recommendations related to the use of differential privacy and disclosure avoidance. For example, the Committee recommended that the Bureau further document the uses of published data, and delay data releases after apportionment in order to allow time for analysis to ensure the privacy of the data. The Bureau plans to formally respond to the recommendations, but as of October 2020, did not provide a timeframe for doing so.

The Bureau still has work left to be done before it implements differential privacy in preparation for delivery of redistricting data by the end of March 2021, and future data products—such as demographic and housing characteristics—that the Bureau plans to produce after the redistricting data. For example, the Bureau plans to hold training sessions regarding differential privacy for its data stewardship executive policy committee this winter. The committee, which is in charge of making privacy policy decisions for the Bureau, plans to make decisions regarding the use of differential privacy on 2020 Census data through early 2021. Additionally, the Bureau still needs to conduct its code peer review on its disclosure avoidance system, which it plans to do before the end of the year.

According to the Chief Scientist, it is likely that certain plans and schedules may need to be updated if the release dates for data products, such as redistricting data, change due to operational impacts from COVID-19. We have ongoing work monitoring the Bureau's progress as it works to implement differential privacy for the 2020 Census.

³¹13 U.S.C. § 9.

³²Differential privacy is a disclosure avoidance technique aimed at limiting statistical disclosure and controlling privacy risk. According to the Bureau, differential privacy provides a way for the Bureau to quantify the level of acceptable privacy risk and mitigate the risk that individuals can be reidentified using the Bureau's data. Reidentification can occur when public data are linked to other external data sources. According to the Bureau, using differential privacy means that publicly available data will include some statistical noise, or data inaccuracies, to protect the privacy of individuals. Differential privacy provides algorithms that allow policy makers to decide the trade-off between data accuracy and privacy.

Agency Comments

We provided a copy of this draft report to the Department of Commerce. The Census Bureau provided technical comments that were incorporated as appropriate.

- - - - -

We are sending copies of this report to the Secretary of the Department of Commerce, the Under Secretary of Economic Affairs, the Director of the Census Bureau, and interested congressional committees. In addition, the report is available at no charge on the GAO website at <http://www.gao.gov>.

If you or your staff have any questions about this report, please contact J. Christopher Mihm at 202-512-6806 or by email at mihmj@gao.gov or Nick Marinos at (202) 512-9342 or by email at marinosn@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made key contributions to this report are listed in enclosure II.



J. Christopher Mihm
Managing Director, Strategic Issues



Nick Marinos
Director, Information Technology and Cybersecurity

Enclosures - 2

List of Requesters

The Honorable Ron Johnson
Chairman
The Honorable Gary C. Peters
Ranking Member
Committee on Homeland Security and Governmental Affairs
United States Senate

The Honorable Carolyn B. Maloney
Chairwoman
The Honorable James Comer
Ranking Member
Committee on Oversight and Reform
House of Representatives

The Honorable Jamie Raskin
Chairman
The Honorable Chip Roy
Ranking Member
Subcommittee on Civil Rights and Civil Liberties
Committee on Oversight and Reform
House of Representatives

The Honorable Gerald E. Connolly
Chairman
The Honorable Jody Hice
Ranking Member
Subcommittee on Government Operations
Committee on Oversight and Reform
House of Representatives

The Honorable Karen Bass
House of Representatives

The Honorable Joaquin Castro
House of Representatives

The Honorable Judy Chu
House of Representatives

The Honorable Steven Horsford
House of Representatives

The Honorable Jim Jordan
House of Representatives

Enclosure I: Court Challenges to the Bureau's 2020 Planned Timeline

Prior to the Coronavirus Disease 2019 (COVID-19) pandemic, the Census Bureau (Bureau) planned to complete data collection by July 31, 2020, and deliver apportionment data to the President by December 31, 2020.

In April, in response to COVID-19, the Bureau announced it would complete data collection by October 31 and deliver apportionment data to the President by April 30, 2021.

On August 3, the Bureau announced it would deliver the apportionment counts to the President by the statutory deadline of December 31, 2020, and it would accelerate completion of its data collection, such that data collection in the field would end by September 30. The Bureau also announced it would accelerate data processing operations in order to meet the statutory deadline. In response to this decision, on August 18, 2020, the National Urban League, as well as other civil rights and civic organizations, local governments, and individuals filed a court case in the United States District Court for the Northern District of California, alleging violation of the Administrative Procedure Act and the Enumeration Clause of the U.S. Constitution.³³

On September 5, the district court issued a temporary restraining order enjoining the Bureau from implementing the acceleration of its data collection and data processing operations, until the Court conducted a hearing on the plaintiff's motion for a stay and preliminary injunction.³⁴

After the hearing, on September 24, the district court issued a preliminary injunction and stayed the Bureau's September 30 and December 31 deadlines.³⁵ The next day, the Department of Justice, on behalf of the Bureau and Department of Commerce, filed an appeal of the injunction and a motion to stay the injunction at the United States Court of Appeals for the Ninth Circuit. On September 28, the Bureau announced that its plan was to conclude data collection operations by October 5, which, according to the Bureau, would still allow it to produce data for apportionment by December 31.

On October 1, the district court, finding that the Bureau had violated the preliminary injunction by both initially continuing the plan to end data collection by September 30 and then by instituting the new plan to end by October 5, issued an order clarifying the preliminary injunction. The district court clarified that, pursuant to the injunction, the Bureau was to use the October 31 deadline from the Bureau's COVID-19 plan for the completion of data collection and, rather than use the December 31 deadline, the Bureau was to use the April 30, 2021, deadline for reporting apportionment data to the President.

On October 7, the Ninth Circuit affirmed the reinstatement of the October 31 deadline, but reversed the injunction as to the December 31, 2020, statutory deadline for reporting apportionment data to the President.³⁶

That same day, the Bureau requested that the U.S. Supreme Court stay the district court's preliminary injunction of the October 31 deadline. The Supreme Court granted the stay on

³³*National Urban League v. Ross*, Case No. 20-cv-05799 (N.D. Cal. Aug. 18, 2020)

³⁴The temporary restraining order also prohibited the Bureau from implementing any actions as a result of the shortened timelines, including but not limited to winding down or altering any census field operations, *National Urban League v. Ross*, Case No. 20-cv-05799 (N.D. Cal. Sept. 5, 2020).

³⁵*National Urban League v. Ross*, Case No. 20-cv-05799 (N.D. Cal. Sept. 24, 2020)

³⁶*National Urban League v. Ross*, Case No. 20-16868 (9th Cir. Oct. 7, 2020)

October 13 pending disposition of the appeal of the preliminary injunction in the Ninth Circuit.³⁷ On October 13, the Bureau stated it would end data collection as of October 15, 2020, and deliver apportionment data on or shortly after the December 31, 2020, statutory deadline.

³⁷*Ross v. National Urban League*, Case No. 20A62, 592 U.S. ____ (2020).

Enclosure II: GAO Contacts and Staff Acknowledgments

GAO Contacts: J. Christopher Mihm, (202) 512-6806 or mihmj@gao.gov

Nick Marinos, (202) 512-9342 or marinosn@gao.gov

Staff Acknowledgments: In addition to the contacts named above, Alexandra Edwards, Jon Ticehurst, Kate Sharkey, (Assistant Directors), Kerstin Meyer, Lisa Hardman (Analysts-in-Charge), Andrea Starosciak, Paulina Maqueda Escamilla, Mark Abraham, Andrew Avery, Michael Bechetti, Anna Blasco, Devin Braun, Alan Daigle, Ahmad Ferguson, Rebecca Eyler, Sam Gaffigan, Amalia Konstas, Ceara Lance, Hannah Laufe, Carlton Maynard, Melissa Melvin, Ty Mitchell, Lisa Pearson, Emmy Rhine Paulé, Rachel Schultz, Andrew Stavisky, Farrah Stone, Peter Verchinski, Shawn Ward, and Tim Wexler made significant contributions to this report.