



Testimony

Before the Committee on Homeland
Security and Governmental Affairs,
U.S. Senate

For Release on Delivery
Expected at 2:30 p.m. ET
Tuesday, March 23, 2021

2020 CENSUS

Bureau Needs to Ensure Transparency over Data Quality, Finalize Plans for Data Protection, and Examine Lessons Learned

Statement of J. Christopher Mihm, Managing Director,
Strategic Issues

and

Nick Marinos, Director, Information Technology and
Cybersecurity

Accessible Version

2020 CENSUS

Bureau Needs to Ensure Transparency over Data Quality, Finalize Plans for Data Protection, and Examine Lessons Learned

Why GAO Did This Study

The Bureau is responsible for conducting a complete and accurate decennial census of the U.S. population. The decennial census is mandated by the Constitution and provides vital data for the nation. For the 2020 Census, the Bureau was required by law to count the population and deliver the counts to the President by December 31, 2020, in order to determine the number of congressional seats apportioned to each state. The Bureau was also required to deliver population counts for redistricting purposes to the states by March 31, 2021.

In response to the COVID-19 pandemic, the Bureau had to make a series of changes to the design and time frames of the census.

GAO was asked to testify about (1) the status of the 2020 Census, (2) ongoing challenges the Bureau faces in completing the count, and (3) lessons learned to inform planning for the next census. To do so, GAO summarized its prior work assessing the operations of the 2020 Census and identifying lessons learned as planning begins for the 2030 decennial census. GAO also included information from its ongoing work examining the Bureau's progress in developing schedules and plans to protect respondent data.

What GAO Recommends

GAO has previously made recommendations aimed at addressing census data quality concerns and using lessons learned from the 2020 Census to inform planning for 2030. The Department of Commerce has agreed with GAO's findings and recommendations.

View [GAO-21-462T](#). For more information, contact J. Christopher Mihm at (202) 512-6806 or mihmj@gao.gov or Nick Marinis at (202) 512-9342 or marinosn@gao.gov.

What GAO Found

The U.S. Census Bureau (Bureau) has critical work remaining to deliver apportionment counts and redistricting data. Specifically, the Bureau did not meet the deadline of delivering apportionment data to the President and announced it would complete the apportionment count on April 30, 2021, because of delays and re-plans in the data collection schedule and changes to the data processing schedule, among other things. COVID-19, executive actions, and uncertainty about the outcomes of ongoing litigation have been contributing factors leading to the delays in delivering apportionment and districting data. As of mid-March 2021, the Bureau was working to complete the remaining processing steps needed to produce the apportionment count, including establishing the final universe of addresses, enumeration statuses, and population count. Regarding redistricting data, the Bureau reported that it is planning to deliver redistricting data to all states by September 30, 2021. Continued focus on the steps needed to complete the apportionment and redistricting files is essential to meeting rescheduled dates.

The Bureau is facing two challenges in completing the count: it has not yet assessed concerns about data quality or finalized plans to protect data.

- GAO has previously noted that late design changes increase risk for a quality census. The Bureau is examining ways to share quality indicators of the census in the near term and is considering how to respond to recommendations made by advisory groups related to, among other things, documenting the quality of the population counts that it provides to the President and the Congress. In December 2020, GAO highlighted key quality indicators that the Bureau should consider providing when it releases the apportionment counts; these quality indicators are consistent with the work of the American Statistical Association and the Census Scientific Advisory Committee. Transparency over what is known and not yet known about census quality would help to increase public and congressional confidence in the quality and completeness of the census.
- As GAO reported in December 2020, the Bureau's original plans and schedules to protect the privacy of respondent data were impacted by the changes to schedules as a result of COVID-19. In February 2021, the Bureau updated its schedule for near-term activities aimed at protecting respondent data in its redistricting data product. However, as of March 2021, the Bureau had not finalized plans related to the protection of respondent data in products expected to be issued after redistricting.

Looking toward the future, it is critical for the Bureau to ensure that it can learn as much as possible about the effects of the design and schedule changes on data collection and quality control operations. The Bureau is updating plans to assess data collection operations and identify resulting lessons learned from the 2020 Census. Actions taken by the Bureau out of necessity to meet compressed deadlines may be considered as options in future censuses. As part of its planning for the 2030 census, it will be important for the Bureau to assess the impact of the 2020 late design changes and the operations' challenges that arose.

March 23, 2021

Chairman Peters, Ranking Member Portman, and Members of the Committee:

We are pleased to be here today to discuss 2020 Decennial Census. Over the past decade, we have issued numerous reports on the 2020 Census. Because of the severity of the challenges and risks that we identified and reported, we designated the 2020 Census as a high-risk area in 2017, and it remained on our most recent High-Risk List issued earlier this month.¹

As you know, the census was taken under extraordinary circumstances. In response to the Coronavirus Disease 2019 (COVID-19) pandemic and related executive branch decisions, the U.S. Census Bureau (Bureau) made a series of late changes to the design of the census. These changes affected the way the Bureau did its work as well as the time it took to do the work. The changes also introduced risks to the quality of data that the Bureau provides for congressional apportionment and redistricting purposes.²

As the Bureau moves toward planning for the next Census in 2030, the agency could better position itself by incorporating lessons learned as a result of this census and addressing open recommendations we have made over the last decade. In this regard, as of March 2021 we have made 113 recommendations related to the 2020 Census, 19 of which have not been fully implemented.

¹GAO, *High-Risk Series: Dedicated Leadership Needed to Address Limited Progress in Most High-Risk Areas*, [GAO-21-119SP](#) (Washington, D.C.: Mar. 2, 2021); *High-Risk Series: Substantial Efforts Needed to Achieve Greater Progress on High-Risk Areas*, [GAO-19-157SP](#) (Washington, D.C.: Mar. 6, 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). GAO'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.

²Census data are used, among other purposes, to apportion the seats of the U.S. House of Representatives and redraw congressional districts in each state.

At your request, our testimony today will describe (1) the status of the 2020 Census, (2) ongoing challenges the Bureau faces in completing the count, and (3) lessons learned to inform planning for the next census.

The information in this statement is based primarily on our recent work assessing the cost and progress of the 2020 Census, identifying quality concerns and quality indicators, and identifying lessons learned as planning begins for 2030.³ To update the status of the 2020 Census, we followed up on our December 2020 work on the cost and progress of the 2020 Census, reviewed recent Bureau announcements about the status of the apportionment count and redistricting file, and interviewed relevant agency officials.⁴

To describe the challenges the Bureau faces in completing the count, we relied on our recently issued reports assessing the operations of the 2020 Census and identifying quality concerns and quality indicators related to Census data.⁵

We also included information from our ongoing work related to the Bureau's plans to protect respondent data by using disclosure avoidance methods. For that work, we collected and reviewed documentation on the status of disclosure avoidance activities, such as milestone schedules. We also interviewed relevant agency officials about their plans to implement disclosure avoidance for 2020 Census data products.

To describe lessons learned, we relied on our report issued and released yesterday that identified challenges and lessons learned as planning begins for 2030.⁶

More detailed information on our objectives, scope, and methodology for our work can be found in the issued reports. We conducted all of the work

³GAO, *Decennial Census: Bureau Should Assess Significant Data Collection Challenges as It Undertakes Planning for 2030*, [GAO-21-365](#) (Washington, D.C.: Mar. 22, 2021); *2020 Census: The Bureau Concluded Field Work but Uncertainty about Data Quality, Accuracy, and Protection Remains*, [GAO-21-206R](#) (Washington, D.C.: Dec. 9, 2020); *2020 Census: Census Bureau Needs to Ensure Transparency over Data Quality*, [GAO-21-262T](#) (Washington, D.C.: Dec. 3, 2020); *2020 Census: Census Bureau Needs to Assess Data Quality Concerns Stemming from Recent Design Changes*, [GAO-21-142](#) (Washington, D.C.: Dec. 3, 2020).

⁴[GAO-21-206R](#).

⁵[GAO-21-206R](#), [GAO-21-262T](#), and [GAO-21-142](#).

⁶[GAO-21-365](#).

on which this statement is based 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

For the 2020 Census, the Bureau was required by law to count the population as of April 1, 2020 (Census Day), and deliver state population counts to the President by December 31, 2020, in order to determine the number of congressional seats apportioned to each state. The Bureau was also required by law to deliver population counts to the states within 1 year of Census Day for redistricting purposes—March 31, 2021, in the case of the 2020 Census.

Even in normal circumstances, conducting the count to meet these deadlines is an enormous and complex undertaking, for which the Bureau spends years planning. This decennial, the COVID-19 pandemic made the effort all the more challenging for the Bureau.

Just over 1 year ago as the Bureau had begun mailing out invitations to participate in the 2020 Census it had to suspend its field operations because of COVID-19 safety concerns. The Bureau then later revised operational timelines in response to the evolving pandemic conditions and Department of Commerce decisions. Figure 1 shows a timeline of these events.

Figure 1: U.S. Census Bureau's Operations Timeline 2020-2021

U.S. Census Bureau's (Bureau) Operations Timeline 2020-2021



Pandemic event denotes an occurrence external to the census that impacted Bureau decisions
 Operational development denotes a key marker or a decision made in implementing 2020 Census operations

Source: Bureau documentation and data from the Centers for Disease Control and Prevention and John Hopkins University of Medicine. | GAO-21-462T

Text of Figure 1: U.S. Census Bureau's Operations Timeline 2020-2021

- Jan. 27, 2020 - The U.S. declares Public Health Emergency for COVID-19
- March 12, 2020 - Bureau begins inviting responses to the census
- March 18, 2020 - The U.S. records roughly 9,000 cumulative COVID-19 cases
- March 18, 2020 - Bureau suspends 2020 Census field operations
- March 28, 2020 – Twenty-five states enact stay-at-home orders
- April 1, 2020 – Census Day
- April 10, 2020 - The U.S. records roughly 500,000 cumulative COVID-19 cases
- April 13, 2020 - U.S. Department of Commerce seeks statutory relief of 120 additional days for delivering data collection
- May 5, 2020 - Bureau resumes verifying addresses and delivering invitations to respond (Update Leave)
- July 15, 2020 - The U.S. records roughly 3.5 million cumulative COVID-19 cases
- July 16, 2020 - Bureau begins phasing in the nonresponse follow-up (NRFU) operation in selected locations
- August 3, 2020 - Secretary of Commerce directs Bureau to implement re-plan of the census to deliver by statutory deadline. Bureau announces compression of data collection and data processing
- August 9, 2020 - Bureau begins NRFU in all remaining areas
- Oct. 13, 2020 - Supreme Court stays a preliminary injunction against the 8/3/20 re-plan, allowing Bureau to proceed
- Oct. 15, 2020 - Bureau ends 2020 Census data collection operations
- Nov. 8, 2020 - The U.S. records roughly 10 million cumulative COVID-19 cases
- Dec. 31, 2020 - Statutory date to deliver apportionment counts to the President of the United States
- Jan. 9, 2021 - The U.S. records roughly 22 million cumulative COVID-19 cases
- Jan. 28, 2021 - Bureau announces that it plans to deliver the apportionment counts to the President by April 30, 2021

- Feb. 12, 2021 - Bureau announces that it will deliver redistricting data to all states by September 30, 2021

As part of its initial COVID-19 response and to promote the safety of its workforce and the public, the Bureau delayed its field operations multiple times. The Bureau and the Department of Commerce also announced plans in April 2020 to delay the delivery of the population counts by 4 months past its statutory deadlines. In May, the Bureau began gradually resuming the nonresponse follow-up operation. Then in August, in a shift of direction, the Bureau announced it would accelerate its time frame, as directed by the Secretary of Commerce, and deliver the counts by the statutory deadlines.

August 3, 2020, re-plan of census. The Bureau developed a nonresponse follow-up “re-plan” that modified certain procedures in order to meet the shortened time frame. To reduce the workload of nonresponse follow-up, the Bureau also made a set of changes to quality control, including reducing the scope of the quality control operation and reducing the number of contact attempts for a quality control interview.

The nonresponse follow-up operation continued under the re-plan for approximately 1 month before the U.S. District Court for the Northern District of California issued a temporary restraining order, followed by a preliminary injunction, which barred the Bureau from operating under the August 3 re-plan.

Due to the injunction, the Bureau ceased re-plan operational changes and the court required that nonresponse follow-up operations continue past the Bureau’s September 30, 2020, deadline. On October 13, 2020, the Supreme Court granted a stay of the injunction, allowing the Bureau to cease data collection, and on October 15, the Bureau ended all data collection operations, including nonresponse follow-up.

Processing census data. In November 2020, Bureau officials stated that they had identified what the Bureau refers to as anomalies as they processed responses for the apportionment count. According to Bureau officials, anomalies during processing are not unexpected, in that they occur with each census, and time is typically built into the schedule to identify and address them.

As of mid-March 2021, the Bureau reported that it had identified 33 anomalies while processing responses for the apportionment count. The Bureau noted that most of the anomalies were coding issues related to

how response data appear and were processed in the Bureau's data files and in the resulting totals. For example, as the Bureau integrated data in group quarter counts, it found mismatches and duplicates. According to Bureau officials, as of March 17, 2020, the Bureau had fixed each anomaly identified.

The Bureau Has Critical Work Remaining to Deliver Apportionment Counts and Redistricting Data

In December 2020, we reported that the Bureau was endeavoring to deliver apportionment data to the President by or as close to the statutory deadline of December 31, 2020, as possible.⁷ However, the Bureau did not meet the statutory deadline. On January 28, 2021, the Bureau announced it would complete the apportionment count on April 30, 2021—or about 4 months after the statutory deadline—because of delays and re-plans in the data collection schedule and changes to the data processing schedule, among other things.⁸ As of mid-March 2021, the Bureau was working to complete the remaining processing steps needed to produce the apportionment count, including establishing the final universe of addresses, enumeration statuses, and population count.

Related to redistricting data, on February 12, 2021, the Bureau reported that it is planning to deliver final redistricting data to the states by September 30, 2021—or about 6 months after the planned date of the end of March 2021. The Bureau reported that delays in redistricting data were due to both the schedule delays for data collection as well as re-prioritization in its data processing schedule. Specifically, as we reported in December 2020, the Bureau re-prioritized tasks needed to produce apportionment counts rather than simultaneously preparing redistricting data. According to the Bureau, this re-prioritization has delayed processing activities necessary to generate the redistricting counts.

In addition, the Bureau reported that it is planning to deliver final redistricting data all at once because of the impact of COVID-19 on the Bureau's data collection and data processing schedule. This is in contrast

⁷[GAO-21-206R](#).

⁸13 U.S.C. § 141(b).

to the Bureau's initial plan to deliver redistricting data state-by-state. According to the Bureau, delivering redistricting data in a single, national delivery will allow them to, among other things, ensure accurate and high-quality data in the least total amount of time to all states.⁹ Before the Bureau can complete the redistricting files, it plans to complete several steps, including ensuring that each record has valid values for major characteristics; applying a disclosure avoidance technique known as differential privacy—described in more detail later—to ensure the confidentiality of responses; and tabulating and reviewing the data for quality control purposes.¹⁰

COVID-19, executive actions, and uncertainty about the outcomes of ongoing litigation have been contributing factors leading to the delays in delivering apportionment and districting data. Continued focus on the steps needed to complete the apportionment and redistricting files is essential to meeting rescheduled dates.

The Bureau Has Not Yet Assessed Concerns about Data Quality and Finalized Plans to Protect Data

Quality Indicators Can Provide the Public with Transparent Information on the Published Counts

We have previously noted that late design changes increase the risks to a quality census. The Bureau made operational changes in response to COVID-19 and subsequent developments that include shortening the time frames for implementing operations, making fewer visits to certain nonresponding households, and reducing the scope and time of quality checks on field data collection and data processing.

In December 2020, we identified a number of concerns about how late changes to the census design could affect data quality. These concerns

⁹On March 15, 2021, the Bureau announced that it plans to provide a legacy-format summary redistricting data file to each state by mid-to-late August 2021.

¹⁰Differential privacy is a disclosure avoidance technique aimed at limiting statistical disclosure and controlling privacy risk. 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.

range from how the altered time frames have affected population counts during field data collection to what effects, if any, compressed and streamlined post-data collection processing of census data may have on the Bureau's ability to detect and fully address processing or other errors. For example, as a result of COVID-19, the Bureau extended its nonresponse follow-up activities through October 15, 2020. Given that the reference date for the census is April 1, individuals who responded to the census may have experienced recall bias, which has the potential to impact an individual's ability to remember where one was living and with how many people the farther removed they are from a specific date in time.

Due to these and other concerns, we recommended that the Bureau ensure its assessments address a number of concerns about how late changes to the census design could affect data quality. The Department of Commerce agreed with our recommendation and described ongoing and planned activities to implement it. The Bureau's activities included planning to release information and findings on data quality as results become available.¹¹

The Bureau is considering how to respond to numerous recommendations made by the American Statistical Association and the Census Scientific Advisory Committee (CSAC) in the fall of 2020, including for the Bureau to document what it knows in near real-time about the quality of the population counts that it provides to the President and to the Congress.¹²

The Bureau also hired a group of independent scientific consultants, known as the JASON group, to assess the Bureau's data processing. In a report the Bureau released in February 2021, the group provided advice on how the Bureau might use various metrics or indicators to assess census quality.

¹¹[GAO-21-142](#).

¹²American Statistical Association 2020 Census Quality Indicators Task Force, *2020 Census Quality Indicators*, (Alexandria, VA: October 2020). The CSAC is an advisory body to the Director of the Census 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 Bureau programs and activities.

In a separate report from December 2020, we highlighted some of the census quality indicators that the Bureau should consider providing when it releases the apportionment counts.¹³ These quality indicators, described in more detail below, are consistent with the work of the American Statistical Association and the CSAC. Disclosing what is known and not yet known about census quality would help increase the transparency and public confidence of the census.

Looking at the extent to which the Bureau used different methods of data collection during the census may provide an indication of the quality of the 2020 Census.¹⁴ While the Bureau believes that self-response from households provides the most accurate data, in an attempt to count every person, the Bureau uses alternative methods when it is unable to collect census data directly from the household, including:

- **Proxy responses for occupied, vacant, and nonexistent housing units:** The Bureau uses proxy responses—information from a neighbor or other knowledgeable person—to collect data when a resident of the household is not available or cannot be found during the nonresponse follow-up operation. For the 2020 Census, preliminary results indicate that, nationally, the Bureau enumerated approximately 7.4 million occupied households using proxies, or 24.1 percent of occupied households in the nonresponse follow-up workload compared to approximately 23.8 percent of occupied households (6.8 million households) in the 2010 Census.
- **Population-count only responses:** Bureau enumerators are not always able to complete the entire census questionnaire for a household and are directed to try to obtain, at a minimum, an indication of the household status (vacant, not a household, or occupied) and if, occupied, the number of people in the housing unit by talking to a proxy. These are referred to as population-count only responses. According to Bureau officials, the number of population-count only responses for the 2020 Census is not yet available and was not reported in 2010.
- **Administrative records responses:** The Bureau incorporated increased use of administrative records into the design of the 2020

¹³[GAO-21-206R](#).

¹⁴In addition to the information presented here on the use of administrative records and proxies in closing nonresponse follow-up cases, the Bureau will likely provide additional information on the use of these methods in the Operational Assessments produced after the 2020 Census.

Census as a major cost saving innovation and to improve the overall quality of the data. Administrative record data refers to information from federal and state governments and third-party vendors that people have already provided, such as information from the United States Postal Service, the Internal Revenue Service, the Social Security Administration, the 2010 Census and the American Community Survey.

These data were used in many new ways in the 2020 Census, such as to identify vacant and non-housing units and to enumerate occupied households if, after one visit, a nonresponse follow-up contact attempt was not successful.

Administrative records were also used in the final phases of the nonresponse follow-up operation. Specifically, households that had received the maximum number of contact attempts (typically six) were reopened, and if the Bureau was still unable to conduct an interview, then it relied on administrative records that met its quality standards. According to Bureau planning documents, while these administrative records used at the end of nonresponse follow-up do not meet the threshold for early removal after the first visit, they do provide adequate data to support apportionment. The Bureau believes its increased use of administrative records helped control cost and improved quality of its data. In total, according to preliminary data, the Bureau used administrative records to close nonresponse follow-up cases for 8.4 million households, or approximately 5.6 percent of households nationwide.¹⁵

- **Count imputation:** To resolve missing household responses following data collection, the Bureau draws data from similar nearby households to determine whether a housing unit exists, whether it is occupied, and, if so, by how many people—a technique it calls count imputation. The Bureau has used some form of imputation since the 1940 Census, and reported on the use of imputation in its operational assessment reports for the 2010 Census.

According to Bureau reporting, in 2010, about 500,000 of 137 million addresses counted in the decennial (0.4 percent) were missing an entire response and the Bureau therefore used count imputation to determine a combination of their residence and occupancy status and

¹⁵The 8.4 million cases includes 8 million occupied households and about 450 thousand vacant households or those determined to be nonexistent.

household size.¹⁶ The Bureau's count imputation in 2010 added about 1.2 million people to the final census count. According to Bureau officials, data are not yet available on count imputation for the 2020 Census.

Nationwide rates provide a high-level indication of overall census quality. However, to provide a more complete assessment and identify outliers, the Bureau also needs to examine the rates at which it uses these alternative methods for specific 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 for this decennial census and are working to provide that data.

The Bureau Has Not Yet Finalized Plans Related to the Protection of Respondent Data

To protect the confidentiality of respondent data as required by statute,¹⁷ the Bureau plans to apply a disclosure avoidance technique, known as differential privacy, to its publicly-released statistical products. These products include redistricting files, detailed race tables, demographic profiles, and demographic and housing characteristics.

In December 2020, we reported that delays and changes to operational schedules as a result of COVID-19 created uncertainty in the plans and schedules related to the implementation of differential privacy.¹⁸ In February 2021, the Bureau updated its plans and schedules for near-term

¹⁶The Bureau imputed whether an address was livable for 0.12 percent of addresses, whether it was occupied or vacant for 0.03 percent, and its household size for 0.24 percent.

¹⁷Title 13 prohibits the Secretary of Commerce, an employee of the Department of Commerce, or local government census liaisons from using information collected for the decennial census for any purpose other than statistical purposes. Additionally, it is prohibited under Title 13 to make a publication that identifies any particular individual or allows anyone other than sworn officers and employees of the Department of Commerce to examine individual reports. 13 U.S.C. § 9(a).

¹⁸[GAO-21-206R](#).

differential privacy activities leading to the development of the redistricting files.

For example, the Bureau plans to publish a new set of demonstration data products—or mock data files to show what the final redistricting files might look like—by the end of April 2021.¹⁹ According to the Bureau's chief scientist, these demonstration products are designed to help data users better understand how the Bureau is using disclosure avoidance for the 2020 Census, and the impact of differential privacy on both the quality of the data as well as the protection of respondent's personally identifiable information.

The Bureau plans to collect user input on those demonstration data products through May 2021. This user input is intended to, among other things, help the Bureau's data stewardship executive policy committee to make key decisions on the implementation of differential privacy and the trade-off between data quality and privacy protections in the Bureau's publicly-available data.²⁰ As of February 2021, these key decisions were expected to be made in June 2021. In addition, the Bureau plans to issue fit-for-use redistricting files—which are expected to be the first data products to use differential privacy protections—by September 30, 2021.²¹

However, as of March 2021, the Bureau had not finalized plans and schedules for the delivery of additional data products, such as demographics and housing characteristics data, to be produced after the redistricting files are delivered. Further, the Bureau had not yet developed time frames for disclosure avoidance-related activities to occur for those data products. For example, Bureau officials in the Directorate for Research and Methodology reported plans to produce demonstration data products for demographics and housing characteristics data, but had not defined a time frame for doing so.

¹⁹The Bureau also published demonstration data products in October 2019, May 2020, September 2020, and November 2020.

²⁰The Bureau's data stewardship executive policy committee is in charge of making privacy policy decisions for the Bureau.

²¹Redistricting products are used, among other things, to delineate voting districts by state.

The Bureau Needs to Examine Challenges and Lessons Learned from the 2020 Census to Inform 2030 Planning

Conducting field data collections amid a global pandemic posed unprecedented challenges to the Bureau. Given the disruptions, it is critical for the Bureau to ensure that it can learn as much as possible about the effects of design and schedule changes on data collection and quality control operations, as well as challenges encountered implementing new methods for collecting data. Moreover, actions taken by the Bureau out of necessity to meet compressed deadlines may be considered as options in future censuses.

In July 2020, for the nonresponse follow-up operation, the Bureau started sending enumerators to count approximately 48.6 million housing units (based on preliminary data) that had not yet responded to the 2020 Census. Nationally, the Bureau reported completing over 99 percent of nonresponse follow-up cases by its October 15, 2020, deadline. The Bureau said technology was one reason it was able to complete the work by its deadline.

Under the August 3 re-plan, when the dates to complete nonresponse follow-up were compressed by 1 month, the Bureau was able to employ other methods to complete nonresponse follow-up cases. Specifically, the Bureau offered enumerators financial awards for working additional hours while completing a minimum number of cases, and removed restrictions on overtime. In some circumstances, the Bureau also offered incentive awards to enumerators willing to travel to particular areas. In March 2021, the Bureau reported \$98.4 million in expenses for enumerator awards, \$795,000 of which was for travel awards.

The Bureau, however, had lower completion percentages ranging between 94 and 99 percent for 10 local geographic areas, in part because of natural disasters and COVID-19. For example, according to the Bureau in Shreveport, Louisiana, short-term closures stemming from a hurricane impacted data collection for 82,863 housing units. As a mitigation strategy, the Bureau shifted the Shreveport operation to telephone enumeration and brought in over 1,200 enumerators from travel teams.

Despite these efforts, the Bureau was unable to complete 22,588 cases in Shreveport before data collection ended. For these cases, the Bureau will

need to rely on alternate methods including imputation, which draws data from similar nearby households to determine whether a housing unit exists, whether it is occupied, and, if so, by how many people.

The Bureau also encountered other issues during nonresponse follow-up, including:

- **Restricted access to large multi-unit buildings.** Enumeration supervisors told us that enumerators were often turned away from accessing multi-unit buildings because of the pandemic. Bureau officials told us that even before COVID-19, access to multi-units had been an issue and that the pandemic made access to multi-unit buildings more challenging, and that they will continue to look for ways to address access issues. If enumerators could not get access to conduct an interview, then the Bureau had to rely on alternative sources of data collection, such as administrative records or proxy data.
- **Inability of supervisors to reassign open cases in a timely fashion.** We found that census field supervisors did not have the authority to reassign cases and had to wait for the field manager to make those reassignments. Bureau officials told us it would consider the reassignment of cases as it moves toward planning for the 2030 Census.
- **Unreliable Census Field Supervisor alerts.** To effectively manage nonresponse follow-up, the Bureau provides data-driven tools for the field supervisors to manage enumerators, including system alerts that identify issues that require the supervisor to follow up. Seventeen out of 40 census field supervisors we spoke to said the alerts were not always useful. For example, census field supervisors told us that alerts would be triggered for enumerators standing too far from the door of the household (inaccurate location) or if an enumerator's start time was outside of the approved time range.

According to Bureau officials, enumerators did not always stand in front of the door because of COVID-19 and this often triggered the inaccurate location alert. Bureau officials also told us that enumerators also logged on early to view their cases for the day and that this would trigger the enumerator start time (too early) alert. According to Bureau officials, they plan to review new technology, which includes field supervisor alerts, as part of its nonresponse follow-up assessment plan.

- **Inconsistent ability to monitor enumerators.** At the end of training, enumerators and census field supervisors are assessed using

automated software that replicates various situations and prompts them for the correct action to take. Enumerators who score less than 70 percent on this assessment are supposed to receive additional on-the-job training led by their field supervisors (referred to as an observation). According to Bureau quality control reports, about 2.45 percent (nearly 8,000) of active enumerators failed the assessment and were not observed, yet continued to work. Bureau officials also told us that they intend to improve the ability to monitor and correct enumerators working (after having failed the initial assessment) without proper training or observation.

- **Dissatisfaction with the technological capability to assign and route cases.** Enumeration supervisors we interviewed mentioned specific frustrations with the technological capability used by the Bureau to assign and route cases, known as the optimizer. Specifically, they determined that the optimizer illogically assigned cases that led to routes that sent enumerators to households in a confusing order, the enumeration start times late in the day, and the uneven distribution of cases among available enumerators.

Bureau officials pointed to the increased productivity for the 2020 nonresponse follow-up operation as an indicator of the optimizer's effectiveness. Officials also clarified that the routes planned by the optimizer can appear counterintuitive, because it calculates the best route and time frame to maximize in-person interviews. However, Bureau officials acknowledged that they might have done a better job of explaining the optimizer during training. Bureau officials said they plan to evaluate the effectiveness of new technological tools such as the optimizer, as well as the training.

- **Late design changes, such as eliminating random reinterviews of nonresponse follow-up cases.** The Bureau conducts a sample of reinterview cases to ensure enumerators follow proper procedures and that results were not falsified.²² Random reinterviews are a selected sample of cases from each enumerator's workload. Because of the August 3 re-plan, the Bureau eliminated the random selection of reinterview cases and reduced the number of contact attempts for conducting a reinterview from six to three. According to Bureau officials, these steps were taken to meet the new time frames for completing nonresponse follow-up, which moved from October 31 to September 30. Bureau officials told us that with the temporary

²²The reinterview operation compares the reinterview data to the original nonresponse follow-up interview to identify enumerator errors and possible falsification.

elimination of random reinterviews before the restraining order, 23.5 percent (nearly 70,000) of enumerators did not have cases selected.

The pandemic also made it difficult to count people living in group quarters, such as skilled-nursing and correctional facilities. The Bureau planned to count individuals living in group quarters between April 2 and June 5, 2020, but revised those dates to July 1, 2020, through September 3, 2020, ultimately collecting information from approximately 272,000 group quarters.

Bureau staff found it challenging to locate a point of contact at some group quarters because they were closed due to the pandemic. Bureau officials told us that in December 2020 they decided to re-contact over 24,000 group quarters to collect data, and that imputation would be used to count individuals at the remaining group quarters still reporting a zero population count.

In addition, group quarter administrators found it difficult to use the eResponse option, an innovation since the 2010 Census, for enumeration. Bureau officials pointed to challenges, including a facility's lack of access or unfamiliarity with Excel, which was the preferred format to complete eResponse. The Bureau accepted eResponse data submitted by group quarters administrators in other formats. Bureau officials acknowledge that additional adjustments and testing will help to improve and simplify this data collection process for group quarters administration prior to 2030.

The Bureau is updating plans to assess data collection operations and identify resulting lessons learned from the 2020 Census. Earlier this month, we recommended that, as part of its planning for 2030, the Bureau evaluate how major operational changes in response to data collection challenges affected the quality and completeness of nonresponse follow-up and group quarter enumerations. At a minimum, this evaluation should address the six key areas previously discussed—procedures for accessing multi-unit buildings, process for the reassignment of cases, supervisory alerts used to manage enumerators, quality control over training assessments, late design changes, and an eResponse option for group quarter data collection. In its comments, the Department of Commerce agreed with our findings and recommendation, and described steps it is taking to address them, including implementing numerous studies to be issued over the next 2 years.

In summary, the Bureau has completed data collection for the 2020 Census. However, critical work remains to ensure the quality, accuracy, and privacy of its final data products. We are committed to continue monitoring the Bureau's efforts to produce high-quality data, and examining the quality and cost implications of the Bureau's COVID-19 response and late design changes. As we begin to look forward to the initial planning for the 2030 Census, we will also report on the Bureau's experience with its innovations to the 2020 Census and any lessons learned that can inform 2030 planning.

Chairman Peters, Ranking Member Portman, and Members of the Committee, this completes our prepared statement. We would be pleased to respond to any questions that you may have.

GAO Contacts and Staff Acknowledgments

If you or your staff have any questions about this testimony, please contact Christopher Mihm at (202) 512-6806 or mihmj@gao.gov or Nick Marinos at (202) 512-9342 or marinosn@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this statement.

GAO staff who made key contributions to this testimony and the underlying reports are Kate Sharkey (Assistant Director), Shawn Ward (Analyst-in-Charge), Mark Abraham, Michael Bechetti, Devin Braun, Ann L. Czapiewski, Alan Daigle, Rebecca Eyler, Ahmad Ferguson, Rob Gebhart, Lisa Hardman, Ceara Lance, Hannah Laufe, Jason T. Lee, Carlton Maynard, Ty Mitchell, Lisa Pearson, Scott Pettis, Andrea Starosciak, Andrew Stavisky, Farrah Stone, Jon Ticehurst, Peter Verchinski, and Timothy Wexler.

This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through our website. Each weekday afternoon, GAO posts on its [website](#) newly released reports, testimony, and correspondence. You can also [subscribe](#) to GAO's email updates to receive notification of newly posted products.

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, <https://www.gao.gov/ordering.htm>.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on [Facebook](#), [Flickr](#), [Twitter](#), and [YouTube](#).
Subscribe to our [RSS Feeds](#) or [Email Updates](#). Listen to our [Podcasts](#).
Visit GAO on the web at <https://www.gao.gov>.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact FraudNet:

Website: <https://www.gao.gov/fraudnet/fraudnet.htm>

Automated answering system: (800) 424-5454 or (202) 512-7700

Congressional Relations

Orice Williams Brown, Managing Director, WilliamsO@gao.gov, (202) 512-4400,
U.S. Government Accountability Office, 441 G Street NW, Room 7125,
Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, DC 20548

Strategic Planning and External Liaison

Stephen J. Sanford, Acting Managing Director, spel@gao.gov, (202) 512-4707
U.S. Government Accountability Office, 441 G Street NW, Room 7814,
Washington, DC 20548



Please Print on Recycled Paper.