

**IN THE UNITED STATES DISTRICT COURT
FOR THE MIDDLE DISTRICT OF ALABAMA
EASTERN DIVISION**

STATE OF ALABAMA, *et al.*,

Plaintiffs,

v.

UNITED STATES DEPARTMENT OF
COMMERCE, *et al.*,

Defendants.

Case No. 3:21-CV-211-RAH-ECM-KCN

DECLARATION OF MICHAEL THIEME

I, Michael Thieme, make the following Declaration pursuant to 28 U.S.C. § 1746, and state that under penalty of perjury the following is true and correct to the best of my knowledge and belief:

1. I am the Assistant Director for Decennial Census Programs, Systems, and Contracts at the U.S. Census Bureau. I have occupied this position since November 2017. The 2020 Census is my third Decennial Census. For the 2010 Census, I was the Chief of the Decennial Systems and Contracts Management Office, providing the primary technology and contract management support for that census. For the 2000 Census, I was the Special Assistant to the Assistant Director for Field Operations working at the national level directing field data collection. In my current role as Assistant Director I am responsible for three Census Bureau divisions: the Decennial Information Technology Division, the Geography Division, and the Decennial Contracts Execution Office. With over 2,000 employees and contractors, these divisions provide all the information technology, geography, and contract management support for the 2020 Census. I am knowledgeable about the progress of the 2020 Census in general and the processing of census data in particular.

2. I am making this Declaration in support of Defendants' Opposition to Alabama's preliminary-injunction motion. All statements in this Declaration are based on my personal knowledge or knowledge obtained in the course of my official duties. In this declaration I:

- Provide background about the progress of the 2020 Census and delays;
- Stress the Census Bureau's commitment to producing high quality, usable, data products from the 2020 Census; and
- Provide background on how the Census Bureau processes data for the 2020 Census and why we are unable to produce redistricting data before the statutory deadline.

Background on the 2020 Census

3. The Census Bureau goes to extraordinary lengths to count everyone living in the country once, only once, and in the right place. The Census Bureau's goal in conducting the decennial census is to count everyone living in the United States, including the 50 states, the District of Columbia, and the territories of Puerto Rico, American Samoa, Commonwealth of the Northern Mariana Islands, Guam, and U.S. Virgin Islands. To that end, we expend significant funds, efforts, and resources in capturing an accurate enumeration of the population, including those who are hard to count.

4. The planning, research, design, development, and execution of a decennial census is a massive undertaking to count over 330 million people across 3.8 million square miles. The 2020 decennial census consisted of 35 operations using 52 separate systems. We monitored and managed the status and progress of the 2020 Census in large part using a master schedule, which has over 27,000 separate lines of census activities. Thousands of staff at Census Bureau headquarters and across the country supported the development and execution of the 2020 census operational design, systems, and procedures. In addition, the 2020 Census required the hiring and management of hundreds of thousands of field staff across the country to manage operations and collect data in support of the decennial census.

5. The complexity and inter-related nature of census operations is echoed in the budget for the 2020 Census. The overall budget estimate for the 2020 Census – covering fiscal years 2012 to 2023 – was \$15.6 billion. The Government Accountability Office (GAO) determined that, as of January 2020, this estimate substantially or fully met GAO's standards and best practices for a reliable cost estimate in terms of credibility, accuracy, completeness, and documentation quality. It is rare for civilian agencies to be so designated, and we are proud that the Census Bureau has achieved this status. As of this writing, the Census Bureau has been appropriated in aggregate just over \$14.2 billion to use for the 2020 Census, covering fiscal years 2012 through 2021.

6. The operational design of the 2020 Census was subjected to repeated and rigorous testing. Given the immense effort required to conduct the census, the importance of the results, and the decade of work by thousands of people that goes into planning and conducting the decennial census, the Census Bureau expends a significant amount of effort to evaluate its planning and design to ensure that its operations will be effective in coming as close as possible to a complete count of everyone living in the United States. Design and testing of the 2020 Census was an iterative process: after each test, we revised our plans and assumptions as necessary.

7. The [2020 Census Operational Plan](#) explains the overall operations of the 2020 Census, including the integration of numerous sub-operations. Further details on most of these sub-operations can be found [on our website](#). A partial list of the major operations for which we have posted detailed operations plans includes:

- a. Local Update of Census Addresses
- b. Address Canvassing
- c. Geographic Delineations
- d. Field Infrastructure and Logistics
- e. Forms Printing and Distribution
- f. Integrated Communications Plan
- g. Count Review
- h. Intended Administrative Data Use
- i. Internet Self-Response
- j. Counting Federally Affiliated Americans Overseas
- k. Non-ID Processing
- l. Update Enumerate
- m. Update Leave
- n. Nonresponse Followup (NRFU)
- o. Response Processing

- p. Formal Privacy Methods
- q. Redistricting Data Program
- r. Post Enumeration Survey (PES)
- s. Count Question Resolution
- t. Data Products and Dissemination
- u. Evaluations and Experiments
- v. Archiving

Census Step 1: Locating Every Household in the United States

8. The first operational step in conducting the 2020 Census was to create a Master Address File (MAF) to represent the universe of addresses and locations to be counted in the 2020 Census. A national repository of geographic data—including addresses, address point locations, streets, boundaries, and imagery—is stored within the Census Bureau’s Master Address File/Topologically Integrated Geographic Encoding and Referencing (MAF/TIGER) System, which provides the foundation for the Census Bureau’s data collection, tabulation, and dissemination activities. It is used to generate the universe of addresses that will be included in a decennial census. Those addresses are then invited to respond, typically through an invitation in the mail. The MAF/TIGER System provides the address and geographic base used by our operational control systems to control responses as they are returned to the Census Bureau. The MAF/TIGER data are used to ensure that each person is tabulated to the correct geographic location as the final 2020 Census population and housing counts are prepared.

9. The Census Bureau continually updated this address list in preparation for the 2020 Census. For the third decade, as mandated by the Census Address List Improvement Act of 1994, the Census Bureau implemented the Local Update of Census Addresses (LUCA) Program to provide tribal, state, and local governments an opportunity to review and update the Census Bureau’s address list for their respective jurisdictions. Between

September 2015 and June 2017, the Census Bureau conducted a 100 percent in-office review of every census block in the nation (11,155,486 blocks).¹ During the in-office review, clerical staff had access to satellite and aerial imagery from federal, state, and local sources, and to publicly available street-level images through Google Street View and Bing StreetSide, which provided the ability to see the fronts of structures, as if standing on the sidewalk.

10. A field operation called In-Field Address Canvassing occurred between August 2019 and October 2019 for approximately 50 million addresses that were not verified in the in-office review. Address Canvassing fieldwork validated roughly 88% of these addresses and the remainder were removed from the universe because the Address Canvassing fieldwork verified that they did not exist, were duplicates, or were non-residential addresses. Some new addresses identified during fieldwork matched addresses already in the MAF as a result of contemporaneous in-office update processes. Other new addresses were added to the MAF.

11. The Census Bureau believes that the Census Bureau's MAF/TIGER System is the most complete and accurate address listing in census history.

Census Step 2: Encouraging Self-Response Throughout the 2020 Census

12. In order to encourage everyone in the United States to self-respond, the Census Bureau designed, tested, and implemented a \$700 million Integrated Communications Program. This included a massive multimedia campaign designed to engage

¹ Statistical geographies establish the geographic areas at which the Census Bureau produces statistics. Census blocks are the smallest geographic areas for which we collect and tabulate data. Census blocks are formed by streets, railroads, bodies of water, and legal boundaries (there are approximately 8 million Census Blocks). Census blocks are aggregated to form block groups, and block groups are aggregated to form census tracts. Census tracts optimally represent about 1,600 housing units and 4,000 people. These statistical geographies nest within governmental unit boundaries, such as municipalities and counties.

stakeholders and partners, and to communicate the importance of the census through paid advertising, public relations, social media content, and the new web site. This was the first census where we made a significant investment in digital advertising, targeting online sites including Facebook, Instagram, paid search engines, display ads, and programmatic advertising.

13. The Census Bureau adapted its outreach strategies in response to delayed census operations due to COVID-19, increasing advertising and outreach to specific areas of the country with lower response rates. We quickly adjusted our messaging, pivoting from our original campaign to encourage people to respond online from the safety of their own homes. The use of micro-targeting allowed the Census Bureau to tailor its messaging, including directing appropriate messages to hard-to-reach communities and those who distrust government, both of which have been traditionally undercounted.

14. The Census Bureau's communications program also relied heavily on partnerships, including with organizations in the State of Alabama. There are two prongs to the Partnership Program, the National Partnership Program that works from Census Bureau headquarters mobilizing national organizations, and the Community Partnership and Engagement Program, that works through the regions at the local level to reach organizations that directly touch their communities. Census partners include national organizations like the National Urban League (NUL), the Mexican American Legal Defense Fund (MALDEF), the National Association of Latino Elected Officials (NALEO), the National Association for the Advancement of Colored People (NAACP), and the U.S. Chambers of Commerce. Major corporations also become census partners. At the local level, partners can be churches, synagogues and mosques, legal aid clinics, grocery stores, universities, colleges, and schools.

Census Step 3: Self-Response

15. The design of the 2020 Census depended on self-response from the American public. In an effort to ensure the most efficient process to enumerate households, the

Census Bureau assigned every block in the United States to one specific type of enumeration area (TEA). The TEA reflects the methodology used to enumerate the households within the block. There were two TEAs where self-response was the primary enumeration methodology: TEA 1 (Self-Response) and TEA 6 (Update Leave). Regardless of enumeration methodology, everyone in the country was able to participate in the census online, by mail, or by phone.

16. TEA 1 used a stratified self-response contact strategy to inform and invite the public to respond to the census, and to remind nonresponding housing units to respond. In total, six mailings including the initial Invitation, reminders, and, if we did not receive a response by the third mailing, questionnaires were to be delivered on a flow basis unless a household responded.

17. Update Leave (TEA 6) was conducted in areas where the majority of the housing units did not have mail delivery to the physical location of the housing unit, or the mail delivery information for the housing unit could not be verified. The purpose of Update Leave was to update the address list and feature data, and to leave a 2020 Census Internet Choice package at every housing unit. The major difference from TEA 1 is that a Census Bureau employee, rather than a postal carrier, delivers the 2020 Census invitation to respond, along with a paper questionnaire. As with other housing units, those in TEA 6 had the option to respond online, by mail, or by phone.

18. Self-response began in March 2020 and was open until October 15, 2020. We are proud to have secured a self-response rate of 67%, higher than the 2010 self-response rate of 66.5%.

Census Step 4: Nonresponse Followup (NRFU) and Quality Control

19. After giving everyone an opportunity to self-respond to the census, census field staff (known as enumerators), attempted to contact nonresponding addresses to determine whether each address was vacant, occupied, or did not exist, and when occupied, to collect census response data. Multiple contact attempts to nonresponding addresses

were done to determine the housing unit status and to collect decennial census response data. This was the Nonresponse Followup operation, or NRFU. Enumerators conducted the NRFU operation using iPhones equipped with “optimization” software that assigned cases based on the enumerator’s availability and to increase efficiency of the operation.

20. In addition to the NRFU operation, the Census Bureau conducted several operations to collect information for individuals who do not live in housing units. The Group Quarters Enumeration collects response information for individuals living in group housing situations, such as college dormitories, prisons, or long term care facilities. The Enumeration at Transitory Locations (ETL) operation collects response information for individuals living at campgrounds and marinas.

21. Cases in the NRFU workload are subject to six contact attempts. The first contact attempt is primarily an in-person attempt. Each contact attempt in the 2020 Census NRFU was either a telephone or an in-person contact attempt (however the vast majority of attempts were in-person).

22. If upon the first contact attempt an enumerator determined an address was occupied and the enumerator was able to obtain a response for the housing unit, then the housing unit was counted, and no follow-up was needed.

23. If upon the first contact attempt, the enumerator was not able to obtain a response, the enumerator was trained to assess whether the location was vacant or unoccupied. Enumerators used clues such as empty buildings with no visible furnishings, or vacant lots, to identify an address as vacant or non-existent.

24. A single determination of a vacant or nonexistent status was not sufficient to remove that address from the NRFU workload; a second confirmation was required. If a knowledgeable person could confirm the enumerator’s assessment, the address was considered vacant or non-existent and no additional contact attempts were needed. A knowledgeable person was someone who knew about the address as it existed on census day or about the persons living at an address on census day. A knowledgeable person

could be someone such as a neighbor, a realtor, a rental agent, or a building manager. This knowledgeable person is known as a proxy respondent.

25. If a knowledgeable person could not be found to confirm the status of vacant or non-existent, use of administrative records could provide confirmation of the enumerator's assessment. The Census Bureau did not rely on a single administrative-records source to determine an address was vacant or non-existent. Rather, multiple sources were necessary to provide the confidence and corroboration before administrative records were considered for use. When used in combination with an enumerator's assessment of vacant or non-existent, corroborated administrative records provided the second confirmation that a nonresponding address was vacant or non-existent.

26. If, upon the first in-person contact attempt, the enumerator believed the address was occupied, but no knowledgeable person was available to complete the enumeration, the Census Bureau used consistent and high-quality administrative records from trusted sources as the response for the household and no further contact was attempted. We consider administrative records to be of high quality if they are corroborated with multiple sources. Examples of high-quality administrative records include Internal Revenue Service Individual Tax Returns, Internal Revenue Service Information Returns, Center for Medicare and Medicaid Statistics Enrollment Database, Social Security Number Identification File, and 2010 Census data.

27. Regardless of whether administrative records were used as a confirmation of vacancy or non-existent status or for the purposes of enumerating an occupied housing unit, the Census Bureau sent, as a final backstop, a final mailing encouraging occupants, if any, to self-respond to the 2020 Census.

28. If a nonresponding housing unit was found to be occupied but no information was gathered on the first attempt, enumerators repeatedly returned. The vast majority of nonresponding addresses in the NRFU workload had the full battery of in-person contact attempts to determine the status of the nonresponding address (vacant,

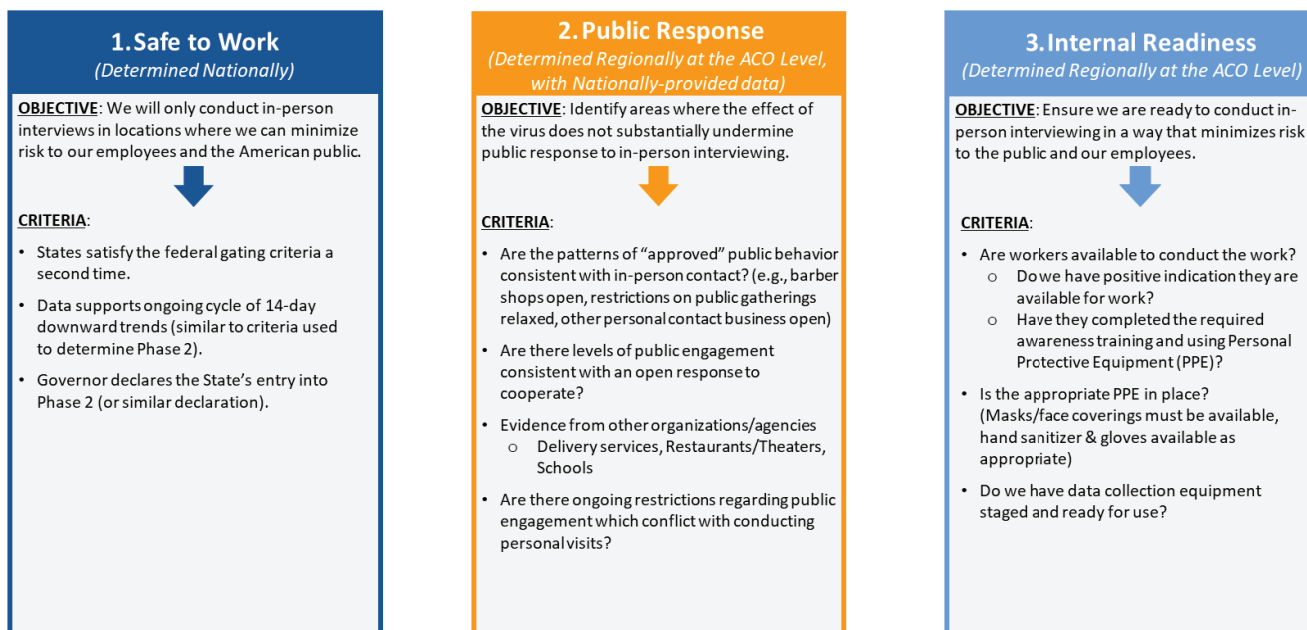
occupied, does not exist) and to collect 2020 Census response data. The full battery of in-person contact attempts also included the ability to collect information from a proxy respondent. Nonresponding units became eligible for a proxy response after three unsuccessful attempts to find residents of a nonresponding address themselves.

29. The Census Bureau arrived at the operational design for NRFU over the course of the decade. Use of administrative records, field management structures, systems, procedures, data collection tools and techniques were proven in tests occurring in 2013, 2014, 2015, 2016, and 2018.

30. While data collection began on schedule, the Census Bureau was forced on March 18, 2020 to announce a suspension of field operations because of the COVID-19 Pandemic. Our original plan was to begin the NRFU operation in most parts of the country in May. But continuing with planned field operations in the spring of 2020 was simply not an option. Many jurisdictions had issued “lockdown” orders. The nation did not know as much about the COVID-19 virus as it does now, and clear public health guidance had not yet been issued. Nor was the Census Bureau able to safely recruit, hire, and train employees for its field operations, and it did not have confidence that households would respond to individuals knocking on their doors seeking responses to the census. Protocols for mask wearing and social distancing were not yet in place and the public health impacts of conducting one of the nation’s largest peacetime mobilizations were unclear.

31. The suspension of field operations and subsequent decisions to adapt field operations were driven by a need to protect the health and safety of the American public; the requirement to implement federal, state, and local regulations on COVID-19; and the desire for a complete and accurate enumeration. We began to re-start operations by resuming our Update Leave operation, resuming pre-NRFU operations in Area Census Offices (ACOs), resuming operations at our paper data capture centers, and resuming fingerprinting and staff onboarding for NRFU workers. The graphic below describes the

criteria we used in our review process for resuming operations during the COVID-19 pandemic.



32. The Census Bureau returned to field operations using a "Soft Launch" approach, meaning that instead of opening all offices at the same time, we instead opened a small number of offices in succession. We opened offices in areas that we believed could be safely started based on COVID risk profiles (developed using CDC, state, and local health guidance), availability of staff, and availability of Personal Protective Equipment (PPE). We needed to acquire PPE, implement social distancing protocols, and work with state and local officials. We opened additional offices throughout the month of July based on detailed daily review of the data about COVID, taking into account state and local stay-at-home orders. We looked for data showing a 14-day downward trend in the area of virus cases, along with sufficient workers to conduct the enumeration, and sufficient available PPE. By August 9 we had begun NRFU in all 248 ACOs. There are 3 ACOs in Alabama. The Census Bureau began NRFU in the Birmingham and Mobile ACOs on August 8, 2020 and commenced operations in the Huntsville ACOs on August 9, 2020.

33. These COVID-19 delays forced the Census Bureau to carry out field operations during hurricane season. Devastating hurricanes in the Gulf Coast area, in particular, limited and slowed the Census Bureau's ability to conduct NRFU operations. Additionally, in large areas of the West Coast, field operations were hampered by conflagrations that caused health alerts due to fire and smoke. And in cities across the country, demonstrations and riots caused further difficulties for in-person enumeration. These challenges forced the Census Bureau to adapt, including by sending travel teams of enumerators to lagging areas.

34. The Census Bureau experienced operational challenges in Alabama in particular. Hurricane Sally hit the state on September 16, 2020 and stopped worked in almost all areas of the state for about 3 days. In some places in Alabama, the Census Bureau was unable to work for two weeks. The Mobile ACO was closed for 5 days, from September 15 to 20, 2020. We also had a higher than average rate of COVID-related closures for our ACOs in Alabama. Our practice was to close an ACO for cleaning when workers tested positive for the virus. The Birmingham ACO was closed 4 times (5 days total), Huntsville was closed 3 times (12 days total), and Mobile was closed 1 time (4 days total). When ACOs were closed, we continued field work and operational management via our use of remote technology, but these closures negatively impacted activities such as hiring, training, and payroll.

35. In light of the COVID-19 delays, on April 13, 2020, the Secretary of Commerce and the Director jointly announced a new Census Schedule and stated that they would seek statutory relief from Congress of 120 additional calendar days. This new schedule set a completion date for field data collection and self-response of October 31, 2020. The proposed schedule called for the delivery of apportionment counts to the President by April 30, 2021 (120 days after the statutory deadline) and redistricting data files to the states no later than July 31, 2021. Congress did not pass such a statute.

36. The Secretary and the Director then announced on August 3, 2020 a “Replan Schedule” designed to meet the Census Bureau’s statutory deadline for reporting apportionment data of December 31, 2020. Litigation ensued, and the Census Bureau was enjoined from attempting to meet the Replan Schedule. After a Supreme Court ruling, the Census Bureau ceased data collection operations on October 15, 2020, having resolved 99.9% of all housing units in the process.

Census Step 5: Post-Data Collection Processing

37. Despite the Census Bureau’s best efforts, the delays caused by COVID, hurricanes and wildfires, along with the normal issues commonly encountered during census processing made it impossible for the Census Bureau to finish apportionment processing and deliver accurate and complete apportionment counts before April 30, 2021 (four months after the December 31, 2020 statutory deadline). However, more important in the context of this declaration is that delivery of apportionment data will occur a month later than the statutory due date for delivering *redistricting* data.

38. The order of phases and operations for processing allows no possible way to deliver redistricting data before apportionment data, as one depends and builds upon the other. Further, the complex processing steps that occur between the apportionment delivery include contingency time for rework if it is required, and the delivery of the complete complement of redistricting products cannot be meaningfully shortened or curtailed without unacceptable risk to the accuracy of the data.

39. Below, I generally describe the post processing operations and schedule. While certain steps in different processing operations may sound similar they are in fact quite different because the steps are iterative. For example, we identify unique persons for the purposes of population count in the early phases, while in the later phases we verify and determine demographic characteristics of every unique person.

A. Incorporate address updates from the field data collection operations into MAF/TIGER

Dates: February 6 – September 27, 2020

40. During the data collection operations, the census field staff can update address, update physical location information, and add addresses. These changes are incorporated into our address and geo-spatial MAF/TIGER databases.

41. Once updated, each address must be associated to the correct state, county, tract, block group and block. Since it is critical for many of our data products (including redistricting data) to associate each address to the correct geography, we verify that the address and geo-spatial updates are incorporated correctly.

B. Produce the Final Collection Address Data Products from MAF/TIGER

Dates: September 27 – Oct 14, 2020

42. Once the benchmark has been created, the final collection geographic data files are produced and verified.

C. Produce and review the Decennial Response File 1 (DRF1)

Dates: October 29 – December 26, 2020

43. The verified final collection geography data are integrated with the response data and we verify our work to ensure accuracy. The next set of activities involves the standardization of the collected information.

44. First, we determine the final classification of each address as either a housing units or a group quarters facility, which is necessary because addresses can change from a housing unit to group quarters and vice versa as a result of field observations. Initial status is set at the start of the data collection operations as either a housing unit or group quarters, but during the enumeration operations, we collect information that informs that classification. For a small number of addresses the classification may change, for example a housing unit may have been turned into a small group home.

45. Next we identify persons in housing unit and group quarters returns. As part of NRFU operation, we reinterview a sample of cases to ensure quality. We incorporate the results of the reinterview for further action in subsequent steps.

46. As part of the Internet self-response option and telephone operation, respondents can provide their data without their Census Identification Number (ID). These cases are assigned an ID which associates them to the final collection geography.

47. We collect data in many ways in the Group Quarters operation, for example on-line, over the phone, on a paper questionnaire, electronic administrative files, and in person using an electronic questionnaire. As a result, we need to standardize the responses across the modes of collection. This step ensures all the data are formatted in the same way for subsequent processing.

48. Finally, for the operations that collect data on a paper questionnaire, some housing units have more people than can fit on one paper questionnaire. The census field staff will use multiple paper questionnaires to enumerate the house. We have to link these continuation forms to form one household.

D. Produce and review the Decennial Response File 2 (DRF2)

Dates: December 26, 2020 – February 26, 2021

49. Once the previous step has been verified, we incorporate the results from the Self-Response Quality Assurance operation that occurred during the data collection phase. This quality assurance operation ensures that data determined to be falsified or incorrect are not added to the Census.

50. As part of the group-quarters operations – a special enumeration procedure used for group housing, such as prisons and colleges – we enumerate domestic violence shelters. Their locations and data are highly sensitive and are handled with special procedures both in the field and in processing. Their data are incorporated at this point in the process due to their sensitivity.

51. Also at this stage in the process, we select a form that will be used as the enumeration of record for those small number of addresses where we receive multiple returns—for example, where one person in a house completes the form online, and another completes the paper questionnaire. This is another important step to avoid duplication, and was particularly important for the 2020 Census, given the multiple modes of response that were offered and the ability to respond without an assigned identification number.

52. As with the prior steps, we continue in the DRF2 file to identify, review, and resolve data anomalies to ensure that the data are accurate.

E. Produce and review the Census Unedited File (CUF)

Dates: February 27, 2021 – March 10, 2021

53. It is in the CUF that we began to incorporate administrative records data as the response data for housing units where we do not have an enumeration but where we have high-quality administrative records data. Incorporating this information thus helps the Census Bureau achieve a more complete census. Administrative record data can include information previously collected by other federal or state agencies, including the Internal Revenue Service and the Social Security Administration.

54. Next, we finalize the status for every housing unit as occupied, vacant or non-existent. Non-existent units are removed from future processing. For every occupied housing unit, we determine the population count.

55. For every housing unit and group quarter, the location is processed by state, county, tract, block group, and block. Then we verify the status (occupied, vacant or non-existent) for every housing unit and group quarter. And in all occupied addresses, we verify the number of persons. For a small number of individuals who live in transitory locations such as campgrounds and marinas, this information will be added later in the process.

56. For unresolved housing units—*i.e.*, those we believe to be occupied but do not have respondent data after all collection operations end—we use a statistical method called count imputation to assign a population count. This was done for less than one half of one percent of housing units in the 2010 Census.

57. The result of these processes is a file that contains records for every housing unit and group quarters along with person records for the people associated with the addresses, although some of the demographic and tenure information may still be missing by this step.

58. As of the date I execute this declaration CUF processing is complete.

59. All of the processes involved in CUF creation are critical predecessors to the subsequent processing operations, including those needed for redistricting.

F. Produce, review and release the Apportionment Counts

Scheduled Dates: March 12 - April 30, 2021

60. On completion of the CUF, we verify and incorporate into the state population counts data that includes the Federally Affiliated Overseas population (such as military personnel serving overseas), the results of the Enumeration of Transitory Locations for each State, and any addresses added late in census data collection.

61. Next, we determine the apportionment counts. Since all housing units and group quarters have a population count linked to a State, we can feed their tabulation into the state-level population counts.

62. To ensure accuracy in the apportionment numbers, the state counts including the overseas population and apportionment numbers are verified by independent teams using different methods. The results of the independent verifications are compared and reconciled, if necessary.

63. We produce the apportionment results using the method of equal proportions. We deliver the apportionment package to the Secretary of Commerce who then delivers them to the President. The President then reports the numbers to Congress.

64. Once again, the Census Bureau does not believe it can deliver apportionment counts sooner than its current working schedule.

G. Produce and Review Census Edited File (CEF)

Scheduled Dates: April 20 – June 24, 2021

65. The next stage is the creation of the Census Edited File (CEF). Whereas the CUF was the basis for the apportionment counts, the CEF provides for the much-more detailed and voluminous data required for redistricting. This process adds any remaining missing values, rectifies conflicting information, and overall ensures a complete set of records for subsequent production of redistricting data.

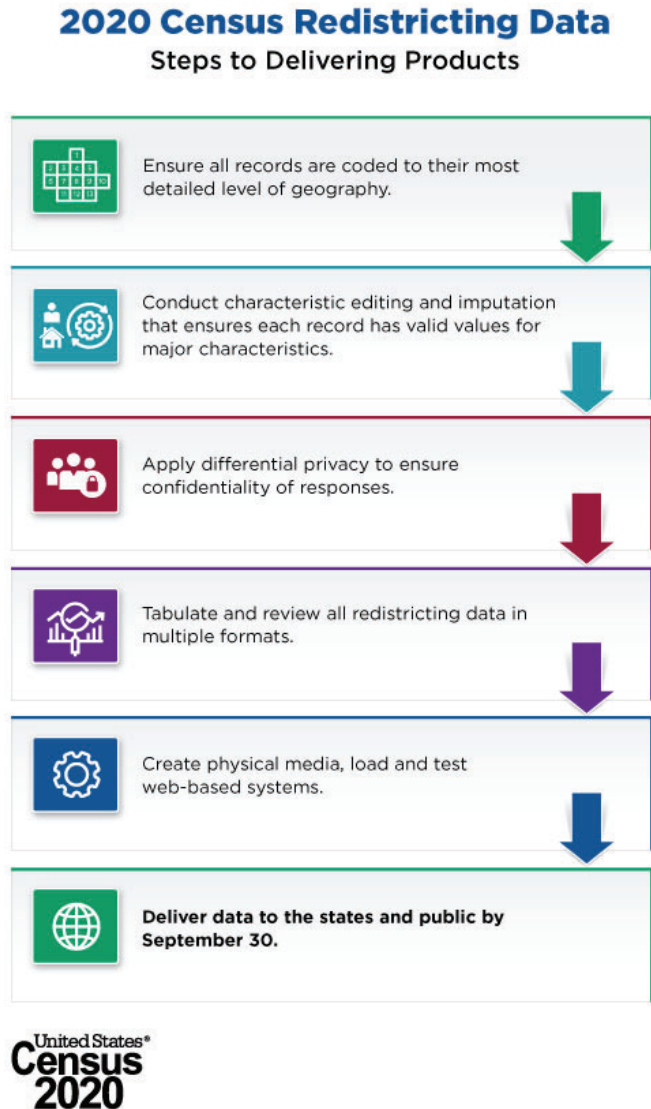
66. It is at this stage that the detailed information about individuals living in the households is compiled (as opposed to the simple population count). This includes complex processing for the race and ethnicity and age information that states must have in order to conduct redistricting.

67. While processing for apportionment only requires accurate population counts, the detailed respondent information needed for the CEF can be conflicting or contradictory and requires application of complex editing rules. Additionally, missing data are accounted for using a statistical process called characteristic imputation.

68. This is a highly complex operation that involves iterative run and review cycles to ensure processing is occurring as designed. And as anomalies are uncovered and corrected, the data must be re-processed.

69. Due to the complexity of the operation, and the iterative cycles of review, it is not possible to accelerate the production of the CEF without unacceptable risks to data quality.

70. The general steps for production of the redistricting data are depicted in the following graphic.



H. Produce and Review Microdata File (MDF)

Scheduled Dates: June 25 – July 18

71. The next stage of the process is to create the privacy-protected Microdata Detail File. At this stage the Census Bureau applies formal privacy protections to prevent

revealing respondent information. This stage is expected to take three weeks; creation of the MDF is not the reason that the Census Bureau will be unable to meet the statutory deadline for delivering redistricting data. In fact the disclosure avoidance procedures completed in the 2010 census processing took 27 days – or nearly four weeks.

72. Application of these protections is not optional. The Census Bureau is required by law to protect the confidentiality of the information provided by respondents. Consistent with that requirement, the Census Bureau designed the 2020 Census Disclosure Avoidance System (DAS) to ensure equal privacy protections for every enumerated person in the country, irrespective of where they live, and that the accuracy of any statistic that we publish will improve as the number of people being measured increases. Our disclosure avoidance approach – differential privacy – accomplishes these objectives by taking detailed privacy-protected measurements of the population at all levels of geography, from the nation down to the individual Census block, and processing these data in descending order from the nation down to the individual block. This method is discussed in greater detail in the declaration of John Abowd.

73. Because of how the method works, the disclosure avoidance algorithm must be applied to the full census data set (i.e., the entire nation) in order to function correctly. Processing these measurements in a top-down fashion, starting at the national level, allows the algorithm to improve the accuracy of the statistics at all geographic levels without impacting the privacy guarantee by leveraging the accuracy of statistics at higher geographic levels. This stage also requires careful review – and, if necessary, re-processing of the data – to ensure the system is functioning as designed.

74. As explained in the declaration of John Abowd, if the Census Bureau were ordered to adopt a different disclosure avoidance methodology at this point in time, it would add significant additional time (at least several months) to the schedule for delivering redistricting data. We would be required to develop an entirely new DAS system, including developing and testing new software. Switching to a new DAS system at the

eleventh hour would also pose significant risks to data accuracy given that the alternative systems (swapping and suppression) are blunt instruments that, unlike differential privacy, cannot be effectively tuned to optimize for data accuracy.

I. Produce and Review Tab file

Scheduled Dates: July 19 - August 16

75. Next, the Census Bureau conducts the tabulation and review of the tabulated census data. Tabular data are easily understandable and usable data tables that the public expects from the Census Bureau. Prior to tabulation, the census data are still in a largely unusable form that would require significant expertise and manipulation on the part of data users to select and understand useful information. Tabulation (literally, formatting and summarizing data into ‘tables’) makes census data easily accessible by state officials or the public. Published tables must be created from the processed data; must be accurate and complete; and must integrate geography, population, and characteristics in myriad combinations.

76. The voluminous and detailed nature of tabulation requires rigorous review and validation that cannot be skipped or shortened without significant risk to the accuracy of these products. This is because tabulation products are created by further processing the base data. The expert review ensures each tabulation product aligns correctly with the base data.

77. As noted in the declaration of James Whitehorne, the Census Bureau announced on March 15, 2021 that it would make a legacy version of redistricting data available to the states in mid-to late- August. While the legacy version requires more data processing expertise on the part of the states, it will be accurate, privacy-protected, and fully usable for redistricting purposes.

78. Finally, during tabulation we also add new fields in the data that make future extraction, summation, and deeper understanding easier for data users to achieve. For example, we may need to add a “voting age” field and update our system so that all

records for people age 18 or over have the that value. This enables “filtering” for the characteristic of “voting age.”

79. If the Census Bureau were to prioritize one State’s legacy version of redistricting data with all planned user tools (to the detriment of the other 49 states), it may be able to deliver a few days earlier than other States, at most.

J. Produce, Load, and Disseminate Redistricting Data

Scheduled Dates: August 17 – September 30

80. The final stage is the review, preparation, loading, and delivery of the official redistricting data. This stage involves the creation of dissemination materials to send the states, the loading of web-based systems, the testing of those materials and systems to ensure they are functioning correctly, and the actual delivery of the data to the states and the public.

81. After we test our materials and systems to ensure that they are functioning correctly, we then create redistricting materials to send to the States. In order to provide convenient access to data users, we also load these data products to our web-based systems. And we deliver the data to the states and the public.

82. While this is happening, we are loading the data.census.gov Data Explorer tool with the entire nation’s data and metadata that allows the system to properly pull and display data. This too must be carefully verified to ensure that data for every geographic level and every table is being properly pulled by the application and displayed. Once all of these materials are ready for release, the physical materials are mailed to the official recipients via overnight mail and the web-tool is made available to the official recipients and the public.

83. If the Census Bureau were to prioritize the DVD/Flash Drive and the data.census.gov webpage for one State’s redistricting data (to the detriment of the other 49 states), it would not be able to deliver the data more than a few weeks earlier than a single national release.

Impossibility of Producing Tabulated Data Prior to September 2021

84. As explained above, the Census Bureau requires approximately five months from the release of the apportionment data on or around April 30 to produce and review the Census Edited File, the Microdata File, the Tab file, and then ultimately produce the redistricting data by September 30. Although the 2020 Census Operational Plan provided for only three months from the planned release of apportionment data of December 31, 2020 to the planned release of redistricting data on March 31, 2021, the Census Bureau now requires an additional two months because of operational changes that the Census Bureau made to expedite the release of the constitutionally required apportionment counts.

85. Specifically, in order to ensure the release of the apportionment counts as quickly as possible, the Census Bureau “decoupled” certain processes that the Census Bureau would have normally completed at the same time. For example, processing of the Census Unedited File (CUF) had pieces set aside to concentrate on processing operations that focused only on population counts for apportionment. This decoupling required designing and creating a second, later CUF format that enabled accurate processing for not only population counts, but also for the demographic characteristics required for redistricting. Under the 2020 Census Operational Plan, the Bureau would have completed this work in parallel with the work on the apportionment counts. This second Census Unedited File would not have been necessary if the Bureau had not, of necessity, decoupled these steps.

86. The delay that has resulted from this “decoupling” of certain processes to prioritize the release of apportionment data accounts for much of the added time in the schedule. The remainder of the additional time accounts for the likelihood that the Bureau will encounter additional “anomalies” that it will need to review and resolve. While it is conceivable that the Bureau may be able to release redistricting data a few weeks earlier if there are zero anomalies, in the Bureau’s experience, this is highly unlikely.

However, if the Bureau is able to produce redistricting data earlier, it will of course do so.

Commitment to Data Quality and Transparency

87. The Census Bureau is committed to unprecedented transparency about the quality of the 2020 Census. The Deputy Director set up a Data Quality Executive Guidance Group (EGG) in April 2020 to evaluate the quality of the 2020 Census. The EGG draws upon expertise of career employees within the Census Bureau in the fields of census operations, statistical methodology, acquisition and use of administrative records, and in the social, economic, and housing subject areas. The group has been meeting continuously since that time and fully supports of the processing steps and the timeline described above to ensure that 2020 Census data are accurate and complete.

88. The Census Bureau is now collecting data for its formal coverage measurement study, the Post-Enumeration Survey. That effort will provide estimates late in 2021 of coverage errors (both undercounts and overcounts) for States and for various demographic groups, similar to what we released after the 2000 and [2010 Censuses](#).

89. In December 2020, the Census Bureau released the 2020 Demographic Analysis (DA) estimates. DA estimates consist of national-level estimates of the population by age, sex, race, and Hispanic origin as of April 1, 2020. These estimates are developed from current and historical vital statistics: birth and death records, estimates of international migration, and Medicare records. The Demographic Analysis estimates are independent from the 2020 Census and are used to calculate net coverage error, one of the two main ways the Census Bureau evaluates the coverage of the census.

90. The 2020 Census has been the most challenging census in modern history. The Census Bureau has faced an unprecedented pandemic, natural disasters, and civil unrest in addition to the already-complicated task of collecting and processing data for nearly 150 million households across 3.8 million square miles. Despite these challenges, the Census Bureau resolved 99.9% of all housing units in the nation and it has planned a


schedule that is designed to achieve the complete and accurate data that will guide the country for the next ten years. At all points, delivery of complete and accurate data has been our overriding priority.

91. I respectfully submit that the intentions of the dedicated professional staff at the Census Bureau will always be to deliver accurate data about the U.S. population and economy, and we continue to do so as we work through the challenges we as a nation have faced over the last year.

I have read the foregoing and it is all true and correct.

DATED and SIGNED:

**MICHAEL
THIEME**

 Digitally signed by MICHAEL
THIEME
Date: 2021.04.11 21:00:34
-04'00'

Michael Thieme

Assistant Director for Decennial Census Programs, Systems, and Contracts

United States Bureau of the Census