March 16, 2021

To: Allison Plyer
            Chair
            Census Scientific Advisory Committee

From: Ron S. Jarmin
            Acting Director
            U.S. Census Bureau

Subject: Partial Recommendations and Comments to the U.S. Census Bureau from the
            Census Scientific Advisory Committee Fall 2020 Meeting

The Census Bureau thanks the Census Scientific Advisory Committee for its recommendations.
We are responding to the partial committee recommendations submitted during its 2020 Fall
Virtual Meeting on September 17-18, 2020.

Your feedback is welcomed to ensure that the Census Bureau continues to provide relevant
and timely statistics used by federal, state, and local governments, as well as business and
industry, in an increasingly technologically oriented society.

Attachment

To: Ron S. Jarmin
            Acting Director
            U.S. Census Bureau

From: Allison Plyer
            Chair
            Census Scientific Advisory Committee

Subject: Partial Recommendations and Comments to the U.S. Census Bureau from the
            Census Scientific Advisory Committee Fall 2020 Meeting

September 17-18, 2020

The Census Scientific Advisory Committee (CSAC) thanks the U.S. Census Bureau for their thorough planning and preparation for this first ever virtual CSAC meeting. The topics covered were timely and salient. The presenters were enthusiastic and engaging. And with very few glitches the technology worked well to support discussion among Census staff and CSAC members participating from remote locations around the country. While in-person meetings support greater information exchange and optimal communication, the virtual platform worked will give the need for physical distancing under pandemic conditions. CSAC thanks the Bureau staff for their extraordinary efforts to make this meeting possible.

Update on the 2020 Census Operations

The Census Scientific Advisory Committee (CSAC) commends the Census Bureau on executing the operations of the 2020 Census in the midst of an unexpected global pandemic. Field work timelines were rewritten from scratch.   Contact attempts were redesigned in a short time so that field work was safe for both Bureau employees as well as for respondents.   Even the communications campaign was adapted, and new advertising generated in some instances.   And the Bureau nimbly executed many advances to plan accelerated post data collection processes. Throughout this process, CSAC has been impressed by the dedication of Census Bureau staff.

After almost a decade of planning, the pandemic outbreak occurred just as field work was starting. In this context, the Census Bureau was able to quickly adapt to changing circumstances and execute data collection in a way that is largely consistent with its planned operational goals. For example, the Bureau was able to meet its goal of having 60.5% of housing units self-respond through mail (as had been done in previous censuses), telephone or internet (two new modes of data collection). Internet self-response was critical to achieving

this goal, particularly given the dynamic of the pandemic. We applaud the Census Bureau on the success of the use of this new technology.

Census methodology requires the Bureau now resolve all remaining housing units through ongoing Non-Response Follow-Up (NRFU), and then complete a series of complex post data collection processes, some of which remain untested and untried.

1) To ensure a successful completion of the 2020 Census in a way that is consistent with its mandate of counting everyone once and in the right place, and based on its scientific and methodological expertise, CSAC recommends that the 2020 Census operational timeline be extended per the Bureau's April 2020 request. Counting everyone once and in the right place, using untested and never-before-used technologies, that must work together with precision, requires time. When the weather isn't right, we postpone the launching of rockets into space. The same should be true of the decennial enumeration, the results of which will impact apportionment, redistricting, funding decisions, legal mandates, and regulatory uses of decennial Census data over the next decade.

   CENSUS BUREAU RESPONSE:

   The Census Bureau does not accept this recommendation for the 2020 Census. Data collection was completed on October 15, 2020.

2) Based on discussions during the Fall 2020 CSAC meeting, the risks to data accuracy from a compressed timeline are substantial. CSAC advises the Census Bureau that the following issues may compromise the accuracy of the 2020 Census under the "Replan" compressed timeline for Non-Response Follow-up:

   - On-going events, including natural disasters and civil unrest as well as pandemic conditions, may make it impossible to complete NRFU by September 30, 2020. This is particularly true in states with the lowest total response rates—particularly Alabama, Mississippi, Louisiana, Georgia, South Carolina, North Carolina, and Florida where weather events may make it impossible for the Census Bureau to complete NRFU operations in some Area Census Offices by September 30, 2020.
   - A shortened NRFU may increase the undercount of newborns as well as other children. Newborns, who are historically undercounted, must be enumerated through household self-response or NRFU and cannot be imputed through tax records. NRFU is an important method by which children under 5 years old are counted.
   - Groups with lower internet access, such as lower income individuals, rural residents, Native Americans, and others are at risk of being more significantly undercounted than in 2010 if NRFU is shortened. These groups are historically undercounted, but the risk of undercounting them is amplified with the pandemic, the reliance on the internet, and the shortened NRFU timeline.

- Reduced contact attempts for self-reported vacant housing units, re-interviews, and self-response quality assurances, without testing the impact on data quality, increases the risk of errors during NRFU.

CENSUS BUREAU RESPONSE:

We understand the Committee's concerns about accuracy of the 2020 Census. The Nonresponse Follow-up operation concluded on October 15, 2020. All states, the District of Columbia, and the Commonwealth of Puerto Rico, achieved total completion rates over 99.0%; and all but one state was at or above 99.9%. While other quality measures for the 2020 Census will be produced over time, it helps to look at self-response rates as an initial metric. Our final self-response rate for the 2020 Census was 67.0%, compared with the final self-response rate for 2010 of 66.5%. Generally, better data comes from self-response, and after a decade of global decline in census and survey participation along with the challenges presented by COVID-19, we were able to beat the 2010 self-response rate.
https://www.census.gov/content/dam/Census/newsroom/press-kits/2020/nrfu-deadline-completion-rates-faq.pdf

3) CSAC remains concerned about the accuracy of the final 2020 Census data based on the shortened time frame for the post data collection processing operation. While previous censuses have required 5 months of post data collection processes, the 2020 Census compressed timeline will only give the Bureau 3 months to complete these tasks, with several data checking processes eliminated. CSAC advises the Census Bureau that the following issues may compromise the accuracy of the 2020 Census under the "Replan" compressed timeline for Post Data Collection Processing:

- Given that large numbers of people are changing their normal residential patterns due to pandemic conditions (e.g., college students, snowbirds), adequate de-duplication procedures for college students, retirees, and others require additional time.

CENSUS BUREAU RESPONSE:

It is possible for college students to be counted by the student housing administrator during the group quarters data collection operation, as well as by administrative records if they lived in off campus housing but went home during the outbreak. Both situations could also result in parents including the student in their housing unit response to the 2020 Census. The Coverage Improvement operation will unduplicate these students if a good address is provided for where they lived while attending school. Regarding the de-duplication procedures in our post-processing systems, they are fully tested and prepared to execute every step of the process. None of the changes in schedule or duration caused by this year's challenges has led to deleting or

altering any processing steps. Our software will run as intended, and we will address all data anomalies and errors, including duplicate responses.

- Elimination of expert review of group quarters by local state demographers through Count Review Event 2 increases the risk that the Census Bureau will publish data with errors in the group quarters population.

CENSUS BUREAU RESPONSE:

FSCPE members were able to provide additional information for group quarter facilities that were not enumerated during group quarters enumeration. These group quarters were contacted again during a modified late group quarters to attempt to collect demographic data for residents or a minimum population count as a last resort.

- Untested post data collection processing systems may fail in ways that the Census Bureau cannot foresee today.

CENSUS BUREAU RESPONSE:

The Census Bureau has thoroughly tested post data collection processing systems, and as stated above, we will address all data anomalies and errors. Additionally, we have been able in this Census to execute pre-processing data checks of states that were approaching completion during data collection. This has given us the ability to discover and address issues earlier. We anticipate that this work will also enable faster and more accurate processing.

In his September 11th statement, the Associate Director for Decennial Census Programs acknowledged that changes to post data collection processing procedures increase the risk of unidentified errors in the collected data.[1] In this CSAC meeting, many of these processes were characterized as redundant. However, redundancy in data checks is necessary to ensure the accuracy of the Census results, just as redundancy in data collection systems was essential to gather the best quality data. Some post data collection processes may seem redundant before executing them, but previous census experience has consistently shown that post data collection activities are an essential tool for ensuring the quality of Census results.

CENSUS BUREAU RESPONSE:

The Census Bureau agrees with this comment. The only redundant process we have eliminated is a review by DSSD of the Master Address File (MAF) 2020 extract We made this decision because there have been decades of work and numerous processes throughout 2020 Census preparations and operations to ensure the MAF is complete and accurate. No other redundant processing steps were eliminated. Any other changes to backend processing only included

---

[1] https://assets.documentcloud.org/documents/7207428/LUPE-Sept-11-2020-Declaration-of-Albert-Fontenot.pdf

reduced calendar time. We mitigated these reductions by working around the clock and adding weekend and holiday production days to the schedule. By closely managing both the process and the risks, our intention continues to be to deliver an accurate, defensible 2020 Census count as close to the statutory deadlines as possible.

4) As a result, the CSAC recommends that the Census Bureau have the time it requested in April 2020 to execute its full battery of data checks to reduce the risk of failing to identify key errors and generate final 2020 Census products that are of comparable quality to previous decennial censuses. Specifically, CSAC believes that the Census Bureau needs the full six months it requested in April 2020 for post data collection processes.

CENSUS BUREAU RESPONSE:

Though the Census Bureau agrees that additional time can further reduce the risk of missing processing errors, we unfortunately have no control over how the U.S. Congress will proceed with legislation to increase the time available for delivery of the 2020 Census results. However, as stated above, by closely managing both the process and the risks, our intention continues to be to deliver an accurate, defensible 2020 Census count as close to the statutory deadlines as possible.

Moreover, the pandemic, natural disasters, late changes to processes, and accelerated timetables are also impacting one of the key tools for measuring decennial Census quality, the Post-Enumeration Survey (PES). Given the known challenges with the planned Post-Enumeration Survey including difficulty of recall and non-response bias, the coverage error in the 2020 Census may not be well measured by the PES. Therefore, it is even more important that the Census execute all originally planned post data collection processes as well as any additional processes envisioned in April 2020 to ensure the 2020 Census data are as accurate as possible.

CENSUS BUREAU RESPONSE:

The Census Bureau accepts the recommendation to conduct a high-quality Post-Enumeration Survey and acknowledges that there are numerous challenges to the planned Post-Enumeration Survey. The Post-Enumeration Survey (PES) is underway. The first two field operations have concluded (Independent Listing and the Initial Housing Unit Follow up) and the third field operation, Person Interviewing, began on September 23.]

In response to these challenges, the Census Bureau formed a PES Quality Team to document challenges and impacts resulting from COVID-19 as well as to identify changes to mitigate and measure errors potential errors. Although COVID-19 has resulted in additional challenges for the PES, such as the potential for more recall error and nonresponse error, we are taking measures to mitigate and correct for many of these potential errors. In general, durations for PES operations are either unchanged or lengthened as a result of the schedule delays.

5) Lastly, to increase quality, CSAC recommends that the Bureau publish daily response rates that include self-response and NRFU completions, at the census tract level. This will support the work of partner organizations in targeting their final outreach efforts to the specific neighborhoods where response rates fall shortest of the 99 percent goal.

   CENSUS BUREAU RESPONSE:

   The Census Bureau does not accept this recommendation for the 2020 Census. The data used to post our NRFU completion rates are not readily available during production at the census tract level. The information we have been providing at the state and ACO levels provides more transparency into the progress of our NRFU operation than in any previous census. We will carry this recommendation forward to our planning for the 2030 Census.

Administrative Records Use

CSAC commends the Bureau for bringing years of developmental work to fruition in the implementation of administrative records (AR) use to reduce the NRFU workload and enumerate a portion of the nonresponding households.

1) CSAC recommends a thorough assessment of this novel application to be presented in a public report. This assessment should include:

   - Estimated cost savings (potential visits saved, and the dollars associated with these visits)
   - Estimated accuracy
   - Model stability over 10 years
   - Ways to improve both the roster building and predictive modeling

   CENSUS BUREAU RESPONSE:

   The Census Bureau accepts this recommendation. The administrative record usage will be documented in the Nonresponse Follow-up operational assessment. The Census Bureau thanks the committee for these suggestions and will attempt to address them in public reports as much as possible.

2) With respect to assessing accuracy, CSAC recommends consideration of the following approaches among others that the Census Bureau may propose:

   - For the set of addresses where prospective AR households were replaced by late self-responses, compare the counts and composition between the two.

- For self-responding addresses that would have qualified as AR Occupied if they had been part of the NRFU workload, compare the counts and composition between the self-responses and the AR Occupied.

CENSUS BUREAU RESPONSE:

The Census Bureau accepts this recommendation and will address these questions in the Nonresponse Follow-up operational assessment.

3) The accurate enumeration of college students both on-campus and off-campus was especially challenging in 2020. Lessons learned from these efforts may prove valuable in the future. CSAC recommends that the Census Bureau conduct a thorough analysis of the accuracy of enumeration of these populations, using whatever methods are available. These should include both the PES and demographic analysis.

CENSUS BUREAU RESPONSE:

The Census Bureau accepts this recommendation. This will include attempts to use the PES to evaluate non-group-quarters coverage only as one limitation will be that group quarters, such as college dorms, were out of scope for that survey.

4) Since young adults have higher mobility rates even without a pandemic, CSAC recommends that the Census Bureau explores (after 2020 Census operations) discussions with the U.S. Department of Education to include an exemption to FERPA[2] to allow colleges and universities to share student information for purposes of the decennial Census enumeration and/or post data collection processing (protected by Title 13 and not usable for other purposes).

CENSUS BUREAU RESPONSE:

The Census Bureau accepts this recommendation.

5) CSAC recommends that the Census Bureau explore a partnership with universities that would facilitate data sharing to improve potential enumeration of students in university locales.
CENSUS BUREAU RESPONSE:

The Census Bureau accepts this recommendation.

6) CSAC recommends that the Census Bureau explore obtaining birth records for the first three months of the year to facilitate roster building.

CENSUS BUREAU RESPONSE:

---

[2] Family Educational Rights and Privacy Act

The Census Bureau accepts the recommendation. The Census Bureau has already obtained the updated birth applications for the first three months of the year from the Social Security Administration. The Census Bureau will evaluate these records to see if there are ways for the 2030 Census that these records could be included.

7) CSAC recommends exploring whether the savings from AR use in NRFU can be applied to improve on-the-ground enumeration.

   CENSUS BUREAU RESPONSE:

   The Census Bureau accepts this recommendation. The plan to use administrative records where they were of high quality for the 2020 Census was part of the plan to improve on-the-ground enumeration. The plan that was researched and tested throughout the decade was that non-responding units with no or lower-quality administrative records could be targeted for relatively more visits as compared to the addresses with higher quality administrative records.

8) CSAC recommends exploring whether data on cell phones, given their ubiquity, can be used to improve counting of the hard-to-count populations.

   CENSUS BUREAU RESPONSE:

   The Census Bureau accepts this recommendation as a possibility to be researched and tested for the 2030 Census.

Differential Privacy

CSAC commends the Bureau for recognizing and demonstrating the vulnerability of classic Disclosure Avoidance techniques. Reconstruction and re-identification risks are serious and are growing with the increase in computational power and availability of auxiliary data sets.

Census data require protection, and CSAC commends the Bureau for its serious commitment to modern and future-proof privacy protection and its development of differential privacy protocols. Further, CSAC notes that the Bureau's implementation of differential privacy at the scale of the 2020 Census via its TopDown Algorithm (TDA) is an exceptional technical achievement. In the course of developing its differential privacy algorithms and code, the Bureau used an exemplary development process, following current best practices and making new contributions to the field.

In addition, the Bureau has made extensive efforts to seek input on use cases from multiple sources, and the compilation of these use cases (especially the collection of Federal Register use cases) is an excellent resource for studying the effects of differential privacy.

CSAC recognizes that the Bureau has made improvements to transparency in its development of differential privacy protections, including its maintenance of a centralized location for updates: 2020 Disclosure Avoidance System Updates. Some aspects of the Bureau's differential privacy efforts are less transparent.

1) CSAC recommends that the Bureau make further efforts to communicate any updates on the decision-making process for the privacy-loss budget and its allocation, and any updates on the timeline for implementation of differential privacy.

    CENSUS BUREAU RESPONSE:

    The Census Bureau accepts this recommendation. We recognize the importance of clear communication of the policy decisions relating to the Disclosure Avoidance System. We will communicate these decisions through our 2020 Census Data Products newsletter, on our Disclosure Avoidance System Updates website, and through the 2020 Census Memorandum Series.

2) While the Bureau has collected many important use cases, CSAC recommends that the Bureau should take substantially more time to catalog methodically the use cases of census data, including funding allocations, legal mandates, and regulatory practices, across all agencies of the federal government as well as at state and local levels.

    This catalog should be publicly available and will help in selecting priority use cases for analysis (see below) and in determining the overall privacy-loss budget and its allocation for the 2020 census. This catalog should be periodically updated going forward to inform decisions about how differential privacy is applied to the American Community Survey, 2030 census, or other census-derived data. Federal-State Cooperative for Population Estimates (FSCPE) members have already begun cataloging state use cases and could be partners in this work.

    In the meantime, CSAC encourages the Bureau to publish the Excel workbook summarizing the use cases collected from the Federal Register on the Census Data Products website.

    CENSUS BUREAU RESPONSE:

    The Census Bureau accepts this recommendation, to the extent that it is feasible within the constraints of the 2020 Census operational schedule. Recognizing the innumerable and diverse downstream uses of census data at all levels of government, in academia, business, and elsewhere, a comprehensive catalog of these use cases would be nearly impossible. Consequently, the Census Bureau has focused on summarizing and cataloging the types of use cases, with a focus on assessing how to measure data accuracy and fitness for use in these diverse contexts. These use cases are cataloged in the 2010 Demonstration Detailed Summary Metrics. As new categories of use cases

Additional rigorous analysis is needed for different use cases, particularly analyses of impacts on funding formulas for federal agencies and Congressional staffers, and analyses of impacts on legal mandates and regulatory practices, including protections for civil rights.

1) CSAC therefore recommends that the Bureau conduct analyses of the impact of differential privacy for priority use cases (funding, legal, and regulatory at all levels of government). An example of such analysis (for redistricting) is the paper "Variability Assessment of Data Treated by the TopDown Algorithm for Redistricting" (Wright and Irimata 2020).

CENSUS BUREAU RESPONSE:

The Census Bureau accepts this recommendation. However, given the diversity and extent of these uses, a comprehensive analysis would not be possible before key decisions on the privacy-loss budget for the 2020 Census would need to be made. Consequently, the use cases represented in the 2010 Demonstration Detailed Summary Metrics are intended to be representative of the broader array of uses of census data. The Census Bureau also believes that these representative use cases, which include a substantial focus on fitness-for-use of the data for the Census Bureau's population estimates program, are highly correlated with key external demographic use cases. The Census Bureau has invested substantial staff resources in performing these analyses but must also rely on our data users to supplement these efforts. Of particular note, we have partnered with a group of experts identified by the Committee on National Statistics to evaluate and improve the ways we are evaluating fitness-for-use of these data. Our public release of the Detailed Summary Metrics and Privacy-Protected Microdata Files are also intended to allow external data users to perform their own analyses to supplement our own internal evaluations. Feedback from these analyses can be submitted to 2020DAS@census.gov. Additionally, the Census Bureau has published a technical assessment of data fitness-for-use for redistricting and Voting Rights Act related uses, which is currently being updated to reflect the Committee's comments.

2) For example, CSAC recommends a careful study of the impact of Differential Privacy (DP) on the Population Estimates program data, which are used for planning purposes and as an input for other data like the American Community Survey. Using the Fall 2019 demonstration data, the differences between DP version and SF1 version of these base data are large.

   CENSUS BUREAU RESPONSE:

   The Census Bureau accepts this recommendation and acknowledges that evaluations of the data's fitness-for-use for the population estimates program and other downstream internal uses are ongoing.

CSAC appreciates the Bureau's efforts in creating the 2010 Demonstration Products, the Sprint II Detailed Summary Metrics and other updates, and the privacy-protected microdata for evaluation by the community of users. Metrics are essential for users to judge the quality and fitness for use of Census data products. The Bureau has developed, computed, and released a set of useful metrics based on the privacy-protected 2010 Census data. These published metrics were instrumental in helping the community of users to recognize problems with the October 2019 release of the 2010 demonstration products. CSAC applauds the Bureau for adapting its algorithms in response to feedback from that community.

3) While the set of published metrics is very useful, CSAC recommends that the Bureau publish further details on some variables (e.g., housing vacancy status - seasonal homes) and that some geographies should be included/better represented (e.g., zip codes, county subdivisions/minor civil divisions). As another example, the Bureau should analyze how aggregating data from small geographic units affects accuracy.

   CENSUS BUREAU RESPONSE:

   The Census Bureau accepts this recommendation. Many of these new disaggregations and new geographic levels are already included in the revised version of the Detailed Summary Metrics (e.g., school districts, MCDs). These additions and revisions are based on feedback we have received from our data users. The Census Bureau welcomes the Committee's suggestions for additional revisions and will conduct an assessment of including them in future versions of the metrics on an ongoing basis.

4) The recommended use case catalog development and rigorous analysis for priority use cases may suggest the need for new metrics, in addition to those metrics that have already been developed. CSAC recommends that the Bureau revisit the list of metrics periodically as the use case catalog and analyses evolve, to see if additional kinds of metrics would be useful.

   CENSUS BUREAU RESPONSE:

The Census Bureau accepts this recommendation. A variety of new use cases, supplied by our data users, are already included in the revised version of the Detailed Summary Metrics. We will continue to expand and update these metrics as new representative use cases are identified.

5) CSAC appreciates the value of the privacy-protected microdata for evaluating data quality but use of these data is challenging even for sophisticated users. To aid further in the assessment of the quality of the privacy-protected data, CSAC recommends that the Bureau release additional versions of the Detailed Summary Metrics, including quality metrics at a finer scale than the current overall means, by releasing means within bins. For example, the current MALPE (Mean Algebraic Percent Error) statistic could be split into the average negative relative error and the average positive relative error, rather than combining the two.   Other statistics might be split at scientifically meaningful thresholds or at variable-specific cut points, like the quintiles of the distribution.

   CENSUS BUREAU RESPONSE:

   The Census Bureau thanks the committee for this recommendation. There is no longer sufficient staff time to incorporate general recommendations into the Detailed Summary Metrics. However, we would consider adding a metric if the committee identified a specific use case not covered by the current metrics. The Census Bureau will continue releasing periodic Privacy-Protected Microdata Files (PPMF) to support external users who wish to expand the metrics.

6) The post-processing within the TopDown Algorithm (TDA) can create positive biases, particularly in small domains where rounding up occurs to avoid negative values. A concern is that these small positive biases can accumulate as small domains are combined to create custom geographies. To facilitate assessment of bias properties for the privacy-protected data, CSAC recommends that the Bureau should release the non-post-processed data used in TDA, which are unbiased estimates with known error distributions.

   CENSUS BUREAU RESPONSE:

   The Census Bureau accepts this recommendation and is currently investigating how to provide this guidance. In order to be a properly supported 2020 Census data product, the Census Bureau will need to develop a set of tools to inform our data users on how to use and interpret the data. We will supply this guidance, though it may not be released with the publication of the official 2020 Census data products and will not be available for the 2010 Demonstration Privacy-Protected Microdata Files. As final decisions are made, we will communicate to the public.

7) The Bureau should make clear what, if any, metrics for 2020 will be computed from 2020 data. The Bureau should make readily available tools for extrapolating from 2010 demonstration metrics to 2020 use cases. A specific suggestion for such a tool is for the Bureau to develop "Generalized Metrics Functions (GMFs)" by analogy to Generalized Variance Functions. A GMF would be obtained by regression of 2010 metrics on 2010 privacy-protected tabular estimates and cell sizes. The fitted regression model could then be used to estimate 2020 metrics, by plugging in 2020 privacy-protected tabular estimates and cell sizes.

   CENSUS BUREAU RESPONSE:

   The Census Bureau accepts this recommendation. We are currently researching how to produce and provide these tools.

CSAC has been asked to advise on prioritization of use cases in the allocation of the privacy-loss budget (PLB) across data products.   Due to the complexities of the disclosure-avoidance system, the implications of the PLB allocation for privacy, for accuracy, and for the privacy-accuracy trade-off are unclear.   CSAC is not aware of either theory or empirical data sets that would offer guidance in addressing these questions.   The Bureau may be required to produce, from administrative records, estimates of undocumented individuals counted in the 2020 Census, for December release with the state apportionment counts, and the Bureau is developing estimates of the number of citizens in each block based on administrative records for CVAP for release in 2021.

8) Given the use cases that CSAC has considered, and the committee's assessment of potentially missing use cases, CSAC recommends that the privacy-loss budget should be prioritized toward the most important use cases in this order:

   - Government funding (federal, state, local)
   - Legal mandates and regulations
   - Community planning (children's & elder services, infrastructure)

   CENSUS BUREAU RESPONSE:

   The Census Bureau thanks the Committee for this recommendation and will take this under advisement. Our primary focus, to date, has been improving accuracy in population totals for political entities, which largely supports the government funding priority that the Committee has identified. Our secondary priority on improving accuracy for the Census Bureau's population estimates program also supports the Committee's proposed priorities. That said, prioritizing accuracy for population and demographic characteristics for arbitrary geographic entities (e.g., the redistricting use case) is challenging because these geographic areas cannot be defined in advance. This is why the Census Bureau has used the principle that the accuracy of any statistic

produced by the DAS must increase as the population in the category or geography increases.

The Bureau may be required to produce, from administrative records, estimates of undocumented individuals counted in the 2020 Census, for December release with the state apportionment counts, and the Bureau is developing estimates of the number of citizens in each block based on administrative records for CVAP for release in 2021. In either case, citizenship status would receive a share of the privacy-loss budget and would reduce the accuracy and usability of other variables.

9) CSAC recommends that if any citizenship variables are part of the December release or CVAP release, the Bureau should assign to these variables a very small part of the privacy-loss budget, such that these data will be more protected. The citizenship data are more sensitive than many other attributes. This attribute is much more correlated within geographic locations, making reconstruction attacks on the data that account for such correlations much more effective in recovering this attribute. The imputations from the Census Unedited File (CUF) to the Census Edited File (CEF) increase the impact of any one person's data on the output, and thus increase the privacy leakage through this attribute. Further, given the inherently large uncertainties in the imputed citizenship attribute, it may be more beneficial to place more of the privacy loss budget on the more accurate tabulations.

CENSUS BUREAU RESPONSE:

The Census Bureau thanks the Committee for this recommendation and will take it under consideration. The Census Bureau's Data Stewardship Executive Policy Committee (DSEP) will make final determinations on the level of privacy protection for these data.

The Bureau's implementation of differential privacy has followed an ambitious timeline under any circumstances, even in the absence of a global pandemic or other challenges. The Bureau is operating under enormous time pressure to make the incredibly consequential and irreversible decision on the privacy-loss budget and its allocation.   But many implications of this decision for privacy, accuracy, and fitness-for-use are currently unknown. The process by which the Bureau will determine the privacy-loss budget allocation is unclear.   Whatever the choice of privacy-loss budget allocation, the Bureau will need to estimate the re-identification risk to ensure sufficient privacy, will need to give users methods for assessing fitness-for-use, and will need to have a backup plan (e.g., allocate some privacy budget) for the future, in case differentially-private data are not fit for some important use cases.   The recommended use case catalog development and rigorous analysis for priority use cases are important for informing how to allocate the privacy-loss budget across uses.

10) CSAC recommends that the Bureau should delay additional releases after the December apportionment release to allow time for these recommended analyses.

CENSUS BUREAU RESPONSE:

The Census Bureau thanks the Committee for this recommendation and will take it under advisement. Public Law 94-171 requires the Census Bureau to produce and deliver the redistricting data to the states, the District of Columbia, and Puerto Rico within one year following Census Day. Therefore, we are focusing our attention to identifying the activities and schedules needed to attempt to meet this requirement.

Planning for the production and release of the remaining 2020 Census data products will restart immediately following completion of the redistricting data planning activities.

Public Comments

CSAC appreciates the Census Bureau enabling public engagement and recommends that the Census Bureau respond in writing to the four written public comments.

CENSUS BUREAU RESPONSE:

Thank you for your recommendations.