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

APR 2 5 2019

MEMORANDUM FOR: Allison Plyer
Chair
Census Scientific Advisory Committee

From: Steven D. Dillingham
Director
U.S. Census Bureau

Subject: U.S. Census Bureau Responses to Census Scientific Advisory Committee Fall 2018 Recommendations

The U.S. Census Bureau thanks the Census Scientific Advisory Committee for its recommendations. We are responding to the committee recommendations submitted during its December 6-7, 2018, meeting.

Your expertise is valuable to ensuring 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

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Fall 2018 Meeting

To: Ron Jarmin

Performing the Non-Exclusive Functions and Duties of the Director

From: Allison Plyer

CSAC Chair

Subject: 2018 CSAC Fall Meeting Recommendations

Date: December 7, 2018

The committee would like to thank the Census Bureau for the well-organized meeting. As always, we are impressed with the expanse of expertise that is represented among the Census staff, and we enjoyed learning about the many important activities underway at the Census Bureau.

The duty of the Census Scientific Advisory Committee is to provide scientific and technical expertise to address Census Bureau objectives and programs. We understand the importance of this work, and we are all most interested in ensuring that our contributions are useful to the work and mission of the Census Bureau. To that end, we encourage the Census presenters to always include a slide with questions at the end of their PowerPoint presentation that they would like the CSAC to address. This will ensure that our comments and recommendations are targeted to those areas where the Census would benefit from our scientific and technical expertise.

I. 2020 Census Update and 2018 End-to-End Test Results

The plans for Census Bureau systems and operations are focused on executing a quality Decennial Census. At this stage in the program, it is important the Census Bureau focus on executing the plans that have been established.

The systems seemed to be working well during the 2018 test and the committee is pleased to see this. The Census Bureau previously estimated that the design of the 2020 Census could result in savings and avoidance of more than $5B compared to 2010. This estimate based on the original prototype analysis seems to be holding up, but CSAC would like an update on the estimated lifecycle budget for the Census and how that relates to the original estimated savings and avoidance.

Recommendations and Comments to the Census Bureau
from the Census Scientific Advisory Committee
Fall 2018 Meeting

CENSUS RESPONSE:
The estimated lifecycle cost of the 2020 Census was updated and released publicly in January 2018. This updated cost estimate totals $15.6 billion. Details of this estimate, including drivers of change from the previous cost estimate from 2015, were released in the 2020 Census Life Cycle Cost Estimate Executive Summary at the link below. Another update of the 2020 Census Life Cycle Cost Estimate will be released in an updated Executive Summary, expected later in the spring of 2019.

https://www2.census.gov/programs-surveys/decennial/2020/program-management/planning-docs/2020-cost-estimate1.pdf

It should be noted that system and process effectiveness cannot be determined by a relatively short PowerPoint presentation. It is recommended the Bureau consider setting up a working group with CSAC composed of Deb Stempowski, Atri Kalluri, Michael Thieme, Jack Levis, Tom Cook, George Ligler, and any other interested CSAC members to do a one-day deep dive into the systems supporting a decennial function starting with NRFU. If the NRFU deep dive provides significant value, the working group can schedule additional reviews of the systems supporting the other 2020 functions.

CENSUS RESPONSE:
The Census Bureau thanks CSAC for this recommendation. Given that we are in the final testing and preparations regarding system readiness and deployment, we regret that we are unable to accommodate this request.

A. The following are recommendations and comments for the 2020 Census:

1. Closely manage change - Change should be closely monitored through a rigorous change control process. Any program change should have its benefit weighed against impact to program cost and schedule. At this late date, it is recommended all "good ideas" be rejected to ensure program stability. Only items essential for a quality census, such as those responding to issues found during the 2018 test, should be approved. "Good ideas" can be logged for future consideration should the available slack remain.

   CENSUS RESPONSE:
   The Change and Risk Management processes are integral parts of the how the 2020 Census Program conducts program and project management. Both processes are regularly reviewed and revised when opportunities for improvement arise. Over the past six months, revisions have included strengthening the following:

   - *Refinements to Risk and Issue Management processes based on feedback obtained from a recent GAO audit results.*
   - *Revisions to Change Management processes to collapse the overarching 2020 Census portfolio process with the systems and contracts processes based on lessons learned from the 2018 End-to-End Test.*

At this time, the Census Bureau is using these processes to make only critical changes necessary for conducting a quality census.

2. Evaluate / measure execution status – Team members should constantly evaluate critical path tasks as well as how much slack may exist in program areas. This should include evaluation of progress using project management tools like Earned Value.

   CENSUS RESPONSE:
   The Census Bureau accepts this recommendation. As you know from our ongoing engagements, conducting a decennial census is no small task. We have established an Integrated Master Schedule (IMS) and a rigorous process for monitoring activities. On a weekly basis, various reports from are produced/issued. One of these reports is an Executive Alert Report that provides the status of 2020 Census operations and systems. In addition, we create weekly Executive Schedule dashboard reports that provide insight into the progress toward achieving the primary goals of the Census. The dashboards reflect the number of "In Progress" activities projected to finish late and the number of future activities projected to start late. The dashboards also provide context around the plan versus actual status variance via identification and analysis of unique issues attributing to the variance and identify key issues for senior leadership awareness or action.

3. Manage risk – A robust risk log should be in place that includes the risk description, probability of occurrence, risk mitigation, and trigger event to execute the mitigation.

   CENSUS RESPONSE:
   The Risk Management process is an integral part of program and project management of the 2020 Census. The process is regularly reviewed and revised when opportunities for improvement arise. Over the past six months, improvements included refinements based on feedback received from a recent GAO audit on Risk Management in the 2020 Census.

4. Automate testing where possible - As much functional, performance and stress testing as possible should be automated. Additional attention should be placed on performance stress testing.

   CENSUS RESPONSE:
   Automated testing is a key performance and quality strategy for the Technical Integrator, who deploys automated testing whenever possible. The goal with test automation is to reduce test time and maximize efficiency by building reusable test processes that increase the quality of the work. The Technical Integrator (TI) Automation Test Team focuses on automating recurring high-priority system test cases identified by the Program-Level Test Team. The 2020 Census Program Test Strategy and TI Automation Test Plan establishes an overall approach for validating the systems under development. The approach is accomplished by ensuring systems are designed, built, and implemented to meet business requirements and are operational upon deployment. The benefits to test automation

include: (1) the ability for the TI Test Team to execute test cases faster and (2) the ability to test complex data combinations and process complex test scenarios to increase test coverage. Test automation also is important in the application of regression test patches to critical system [i.e., Internet Self Response (ISR)] enabling reduced test times and allowing patches to be deployed to production quicker.

Performance, scalability, and stress testing is another focus area for the TI test team. TI creates and updates performances test scripts for integration and business thread end-to-end testing. This allows the team to efficiently performance-test systems in a repeatable way and also to stress test them to ensure that the system can handle the peak loads during 2020 Census Operation.

5. Determine and publish success metrics – As the goal is to have a quality census while avoiding more than $5B in cost, metrics and methods on how to measure and communicate this result should be determined. This should be done well ahead of the beginning of the Census.

   CENSUS RESPONSE:
   We accept this recommendation. As the goal is to have a quality census while avoiding more than $5B in cost, metrics and methods on how to measure and communicate this result should be determined. This should be done well ahead of the beginning of the Census. Each operation released a Detailed Operational Plan (DOP) to the public. The DOPs outline operational success measures and provide the input requirements for the Performance Measurement reports used to track and monitor cost and progress for the 2020 operations.

6. Success metrics should also include operations metrics that are leading indicators to guide efficient field execution.

   CENSUS RESPONSE:
   We accept this recommendation. Success metrics should also include operations metrics that are leading indicators to guide efficient field execution.

   The Bureau's 2020 Census success metrics include operational metrics to guide efficient field execution. Each field operation has a series of reports which cover progress, performance, quality control, and each stage of the specific field process. Those metrics track data by geography where necessary to inform field and Decennial management on how to adjust a given strategy for the highest amount of efficiency.

The types of data we track include: case workloads, assignments, status, and completion (versus planned progress); employee availability; actual cost (versus expected budget); cases selected for quality control, quality control case status, and quality control case completion; addresses completed per hour, miles charged per address, and hours worked per day.

7. Similarly, the Bureau should plan how it will measure quality.

   CENSUS RESPONSE:
   The Census Bureau will measure the coverage of the 2020 Census through its coverage measurement program. This program includes demographic analysis, as well as a post-enumeration survey. We will also evaluate many of the census operations through our program of experiments, evaluations, and assessments. More information about these programs can be found on our website.

8. Proper communication of both quality and efficiency results will be important.

   CENSUS RESPONSE:
   As in past decennial censuses, we will monitor undercounts of specific populations and use our robust communications and partnership program to communicate those results to the public.

9. Consulting staff from Statistics Canada who have significant experience in internet mode of collection may provide insights into how to improve self-response rate.

   CENSUS RESPONSE:
   We accept this recommendation. We have had ongoing communications with Statistics Canada about various collaborations and shared methodological challenges. Statistics Canada recently provided several documents about their plans to maximize internet response, which have been shared between the Decennial (ADDC) and Research and Methodology (ADRM) directorates.

B. The following are recommendations regarding the 2018 End to End Test:

1. CSAC requests a detailed report on the use of internet self-response, including trouble-spots or challenges that respondents faced. This should include analytics based on logging users' experience, such as where they dropped off on the form.

CENSUS RESPONSE:
We appreciate your interest in the 2018 End-to-End Census Test results. The study plan for the evaluation of ISR in the 2018 End-to-End Census Test included analytics such as those mentioned in your recommendation. The resulting report, the 2018 Census Test Operational Assessment for ISR, should be in development this summer and the results should be available by the fall.

2. CSAC remains concerned in the usability of the online form and asks that interested members are given access to a demonstration instance of the form.

   CENSUS RESPONSE:
   The Census Bureau appreciates the committee's interest in the ISR instrument. The ISR instrument has been and will continue to be tested extensively by our usability staff as well as our subject matter experts and operational components. At this time, no demonstration instance of the form is available for use. Once we fully develop the 2020 Census ISR instrument, we would be happy to demonstrate it to CSAC.

3. The committee is very impressed with the improvement in the cases completed per hour during NRFU. CSAC would like information so we can better understand the items that contributed to the improvement.

   CENSUS RESPONSE:
   The Census Bureau thanks the committee for the positive feedback and can provide more information on the increase in productivity rates. The productivity rates for NRFU have been influenced by many factors, but the primary drivers are the move to an automated data collection instrument, the use of the MOJO system to optimize the work for the enumerators, and the use of administrative records to reduce the number of contacts for NRFU cases. Use of an automated data collection instrument reduces the time that enumerators spent with the completion and custody of paper questionnaires in previous censuses. The MOJO system ensures that NRFU cases get assigned to enumerators in the most efficient way and guides the enumerators in completing their workload in the optimal order and at the best time of day. Using administrative records to identify vacant and delete housing units, as well as to enumerate households, has reduced the number of contact attempts required to complete NRFU cases. All of these factors have made the enumerators more efficient, leading to higher productivity rates.

II. Cybersecurity Update

1. It's difficult to assess or provide feedback on the cybersecurity approach without an explicit articulation of the threat model under which the Census is operating. That detail is perhaps simply not shareable with this group, but ideally the Bureau should frame this conversation in terms of the answers to the question: against what types of entities (with what motivations)

are we worried about defending the Census infrastructure, data, and public perception? The threat model is also useful background for red teams, penetration testers, and bug bounty participants.

CENSUS RESPONSE:
The Census Bureau is working closely with various federal agencies and commercial partners in identifying and protecting against key threat areas that encompass systems, services, and operations. We also are strategically designing redundancy and contingency combined with recoverability and availability of our systems, services and operations. The Census Bureau also implemented a vigorous test plan that ensures state of the art techniques that are used to test the ability of our systems and services to stand up in the ever-changing methods of attempts to compromise them.

2. CSAC is interested in receiving more detail on how non-ID response is implemented and how the potential for exploitation has been mitigated. On the surface, issues include impersonation of households in key populations, and overwhelming the deduplication infrastructure to make it difficult to assess the "true" submission for a household.

   CENSUS RESPONSE:
   The addresses for all non-ID responses are compared with records in the Census Bureau's address frame. In the event of a match, the response data is associated with that Census address record, and the address is removed from the nonresponse follow-up universe. In the event of no match, but successful assignment of the non-ID respondent address to a census block, the address is sent to a verification process. This verification can be performed clerically, using available online geographic reference sources. Failing to verify a response using online sources, the Census Bureau will perform verification through fieldwork. Non-ID responses are subject to the same quality assurance efforts applied to all of 2020 self-response. Below is a summary of that operation:

   - *Self-Response Quality Assurance (SRQA) will examine 2020 Census self-response data collected by the internet questionnaire, telephone interviews conducted by Census Questionnaire Assistance staff, or paper questionnaires returned to the Census Bureau. This will include responses associated with a Census identification (ID) number, or those without an ID, which are commonly referred to as "non-ID cases."*
   - *The SRQA system includes automated checks, as well as analysis and/or field follow-up for responses identified as higher risk for potential falsification.*
   - *SRQA analysts will also use Business Intelligence (BI) tools to analyze the response data.*
   - *After reviewing the responses identified by the automated scoring mechanisms as potentially falsified, SRQA analysts will review the cases and determine which ones require field follow-up. Census field staff will attempt to re-collect the response data for selected cases by traveling to the address associated with the original response and conducting an interview at the door.*

- *Each case sent for field follow-up will be assigned a final disposition, based on the outcome of the fieldwork. Business rules will determine whether responses should be included or excluded from the final counts.*

The deduplication of census responses is a separate process that occurs after the conclusion of response data collection. The Census Bureau does not anticipate Non-ID response processing will overwhelm this process.

3. It seems plausible that key individuals working on the Census may be subjected to attacks more sophisticated than basic phishing schemes. CSAC would like information on how the Census is minimizing their phishing surface area and mitigating advanced, perhaps persistent, attacks on key individuals. As a strawman example: Are key contractors disclosing their Census affiliation on LinkedIn a risk that should be mitigated by deferring that disclosure until after their assignment is complete?

   CENSUS RESPONSE:
   The Census Bureau is aware of the methods that can be used as part of the phishing schemes. We are working with industry leaders and various federal agencies to implement strategies that not only identify potential phishing attacks, but also prevent such attempts. While technical solutions that minimize the effect on system of systems and operations of such phishing attacks are in place or being put in place, the Census Bureau acknowledges the risk associated with sophisticated phishing schemes, and is diligently working to further improve contingency plans.

III. Disclosure Avoidance for Block Level Data and Protection of Confidentiality in Public Tabulations

The CSAC thanks John Abowd for his presentation on Disclosure Avoidance, which helped several members of the committee get an understanding of database reconstruction, and the work the Bureau has been doing to modernize disclosure avoidance.

Recommendations:

1. Concentrated Differential Privacy (CDP) or Renyi Differential Privacy (RDP) allow for a more refined privacy analysis than pure differential privacy. CSAC recommends that the Census Bureau consider reporting the CDP/RDP parameters, in addition to the pure Differential Privacy parameter.

   CENSUS RESPONSE:
   The Census Bureau acknowledges this recommendation and will consider it for future disclosure avoidance systems. We agree that Concentrated Differential Privacy (CDP) and Renyi Differential Privacy (RDP) are important contributions to the differential privacy toolbox.

RDP and CDP give privacy guarantees that can be interpreted in terms of approximate (ε,δ)-differential privacy and not pure differential privacy. Analyzing an algorithm under RDP/CDP usually gives a tighter privacy analysis in two cases: (a) when the noise added follows a Gaussian distribution (and not Laplace or Geometric), then there is a big difference between the guarantee under approximate DP and the guarantee under RDP, and (b) when m differentially private algorithms are run in sequence, then the privacy guarantee grows as a multiple of $\sqrt{m}$ rather than as a multiple of m (as in pure DP).

We have elected not to use RDP/CDP for the following reasons:

- *The mathematics for pure ε-DP are better developed than for any of the approximate (ε,δ)-DP methods.*
- *We have found it easier to explain ε-DP to stakeholders than (ε,δ)-DP.*
- *We have a better understanding of the legal implications of ε-DP than of any of the approximate (ε,δ)-DP approaches.*
- *We use Geometric noise to achieve a pure DP guarantee and to avoid floating-point-based side-channel attacks. Gaussian noise does not give a pure DP guarantee. Analyzing Geometric or Laplace noise under RDP/CDP does not result in significantly tighter privacy analysis (as illustrated in the RDP paper [1], and our own experiments in other projects).*
- *Our current algorithms have a small number (the number of levels in the geographical hierarchy) of rounds of sequential noise addition steps, limiting the noise reduction from use of RDP/CDP.*
- *For the 2020 Census use case, we have worked out semantics for pure ε-DP that incorporate our use of "invariants," quantities that are to be published without noise. We are beginning to understand the impact of invariants on traditional DP, but we have not considered the impact of invariants on either CDP or RDP.*

Realistically, at this point in the development of the disclosure avoidance system for the 2020 Census, it would be quite difficult to incorporate RDP or CDP. However, it may be appropriate to review them for a future data product, such as the American Community Survey or the Economic Census. [1] https://arxiv.org/abs/1702.07476

2. The algorithms that were outlined in the presentation involve several "hyperparameters" such as the split of the privacy budget amongst the various levels of the hierarchy. These hyperparameters could be tuned to optimize a desired benchmark. CSAC recommends that the Bureau discuss at the next meeting what benchmarks it plans to optimize in this selection.

   CENSUS RESPONSE:
   The Census Bureau accepts this recommendation. We currently expose the ability to tune several hyperparameters in the 2020 Disclosure Avoidance System, including the split of the privacy-loss

budget between the various levels of the geographic hierarchy, weighting of the workload tabulations, use of the privacy-loss budget for computing differentially-private estimates of error, and many other internal algorithmic choices, as well.

We have engineered the ability to tune many of these hyperparameters. Weighting of workload tabulations is exposed as part of our implementation of the High-Dimensional Matrix Mechanism (HDMM), for example, which serves as the primary driver of our choice of differentially private measurements to take.

To date, we have run a number of experiments investigating the effects of varying these hyperparameters, but considerable additional testing will be required, and discussion of testing strategies is welcome. In addition to discussion of tuning-experiment methodology, we would appreciate any additional feedback on benchmarks to use in optimizing hyperparameters.

We have evaluated error (and, hence, tuned algorithm performance to improve) or in the future intend to evaluate error, in the following ways:

- *using on the L1 and L2 norms (e.g., mean absolute error versus root-mean square error)*
- *using absolute and relative (i.e., percentage) error*
- *examining error in the frequency of zeros (sparsity measures)*
- *iterating error metrics by table (e.g., for each of the tables reported in PL94-171)*
- *examining error on detailed counts (full crosses of all relevant variables) versus error on marginal counts*
- *investigating on-average error, variance of error, and frequencies/distributions of error*
- *studying error for the extant geographic boundaries as drawn in 2010, for standard hierarchy geographic units (nation, state, county, ..., block)*
- *expanding the geographic hierarchy error measures to include some non-standard-hierarchy geographic units*
- *expanding the geographic hierarchy analyses to study error in classes of anticipated 2020 geographic units.*

We welcome help, especially, with:

- *suggestions for error metrics that we should consider but have not*
- *suggestions for how to pare the multidimensional error metrics down to a more streamlined set for final, executive consideration*
- *detailed use-cases motivating particular metrics.*

As part of this work, we have sought the advice of both external and internal stakeholders to identify common and important use-cases for the decennial census data that might suggest natural metrics. Any further help identifying use-cases of this sort would be most welcome.

3. The question of how differential privacy methods would affect the microdata products was not entirely addressed. The ACS microdata have a wide range of stakeholders, and we recommend the Bureau actively engage with the user community to explain and discuss the implications of applying differential privacy to those products, including the effect on data quality and particular use cases.

   CENSUS RESPONSE:
   The Census Bureau accepts this recommendation. We understand that the CSAC is concerned with the impact of differential privacy on both the decennial census and on the American Community Survey (ACS), and specifically on differential privacy's implications for microdata. To date, no decisions have been made regarding the use of differential privacy in the ACS, and all work on the use of differential privacy in the ACS is tentative and exploratory. We believe that any discussion about the impact of differential privacy on the ACS is premature, but that it is helpful to start the discussion on how disclosure avoidance techniques have always impacted the data in the past and that differential privacy is an evolution of that process.

   The error measures we reported in the prior CSAC meeting were limited to L1 error measures evaluated on decennial census workloads/data. As mentioned in the reply to recommendation 2 in this section, we are currently working on evaluating and quantifying the distribution of error relevant to a number of uses of the underlying data; for example, redistricting is a particularly notable use-case, and we are attending specifically to the distribution of errors relevant to it.

   We welcome additional feedback on use cases that can be transformed into evaluation criteria. When provided with such use cases, we will apply them both to work on the decennial census and to our exploratory ACS work.

   During our work with the decennial census, we have found that it is quite technically challenging to design statistically efficient, tractable algorithms that generate differentially private microdata and that respect a list of complex invariants. Indeed, such invariants are often a primary factor driving technical challenges. However, the Census Bureau has done the original science on incorporating invariants into differentially private publication systems. This science strongly suggests that virtually all invariants should be eliminated. The invariants are substituting for a reasoned analysis of particular use cases by imposing arbitrary "accuracy" standards on the publications. These standards do not have any basis in law, and therefore do not deserve any automatic designation as an invariant.

   Nevertheless, we intend the methods we are designing to generate formally private microdata, and our work to date has been in pursuit of this goal. The successful generation of microdata that implemented only the invariants approved by the Data Stewardship Executive Policy Committee (DSEP) for the 2018 End-to-End test demonstrates that these methods are workable.

Although we plan to generate similar formally private microdata for the production run of the 2020 Decennial Census, it is important to note, as some members of our team have in publications elsewhere, that formally private microdata are not suitable for arbitrary analyses. Tabulations that are not in the workload used by our algorithms and for which we do not explicitly expend privacy-loss budget to estimate error may in some cases be reasonably accurate (as a result of correlation with other queries, on which privacy-loss budget was expended), but without expending privacy-loss budget to improve accuracy on a query or to estimate its error, we cannot make strong statements about its accuracy. This property – the inability to guarantee low error on arbitrary tabulations – is a fundamental property of all disclosure avoidance systems, and is not specific to differential privacy. That said, the query set supported by the 2020 Disclosure Avoidance System is substantially larger than the query set supported by the PL94-171 and SF1 tabulations that were produced for the 2010 Census because the fully saturated contingency table gets allocated privacy-loss budget at each level of the geographic hierarchy. This means that many queries not optimized in the workload are, nevertheless, fully analytically valid. Their margins of error may be used to assess fitness for use. An example of an analytically valid query using the 2020 DAS for the person universe would be any five-way cross tabulation (age x sex x race x ethnicity x relationship to householder) at any level of geography, including block. Such queries were not analytically valid in the 1990, 2000, and 2010 Census SF1 products. They could only be analyzed from the public-use microdata sample, where the geographic area was limited to public-use microdata areas.

For example, ACS public microdata has in the past been the subject of criticism when disclosure-avoidance techniques distorted the age distribution of persons aged 65 years and older [3]. Standard swapping methods alter inferences based on some published tabulations, and suppression-based disclosure-avoidance methods create non-ignorable missing data patterns in the published data. These issues of analytical validity are not new with the use of differential privacy. What is new that differential privacy makes the implications of disclosure-avoidance noise for inference based on tabulation queries explicit and transparent. This property forces us to make clear decisions about trade-offs in both privacy loss versus accuracy and in the relative priority (in terms of HDMM weight or privacy-loss budget allocation, for example) given to competing workload tabulations.

The foregoing discussion is not specific to microdata: its effects may be most obvious to users when browsing microdata, but the same concerns apply when using disclosure-avoidance techniques to produce tabular data releases as well. There is at least one class of concerns specific to microdata, however: complex post-processing techniques are required to convert noisy DP measurements into microdata, especially in the presence of complex invariants. This post-processing typically improves accuracy, but at the same time makes noise distributions difficult to characterize a priority; as a result, sophisticated techniques or explicit expenditure of privacy-loss budget are needed to get high-quality estimates of error for tabulations calculated on the resulting microdata. This issue is worth further consideration, and we welcome feedback on it.

As a relevant aside, we note that the theory of differential privacy does not preclude release of the simpler, noisy DP measurements used to generate the microdata. Unlike the final microdata, these measurements are not integer-valued, may be negative, and are not consistent with one another; however, the measurements tend to come with clear, simpler error distributions or estimates. Therefore, these measurements could potentially serve as a supporting source of information for sophisticated users interested in the impact of disclosure-avoidance noise on statistical inference, if these measurements were approved for release alongside the official microdata. ([3] https://www.frbsf.org/economic-research/files/wp10-03bk.pdf).

4. In anticipation of the release of Decennial Census products, the Census Bureau should begin to develop and test explanations for non-technical users about the new application of differential privacy and how this may or may not affect different use cases.

   CENSUS RESPONSE:
   The Census Bureau accepts this recommendation. We agree that this is a complicated subject and difficult to explain to nontechnical users. The best understanding will come from seeing actual data tables using prototype data. The first opportunity will be the release of the 2018 Census Test prototype redistricting data scheduled for March 28. We will conduct webinars and briefings to help data users better understand how the new methods were supplied and how to understand the coding and source material behind the data. We are also supplying additional test files in the weeks that follow and additional instruction. Traditional methods added noise at an undisclosed rate with sometimes negative (but not apparent to the user) consequences. The new method gives the user the tools needed to assess whether the protected data serves the user's need for precision.

IV. Reaching Hard to Count Populations Update

The committee believes that improving the response rates from traditionally hard-to-count populations should be a focus of the 2020 Census team and appreciate the thinking so far about how the main components of the 2020 Census can contribute to this goal.

1. The committee is interested in a report on the analysis of the end-to-end test. In particular, we would like to know the self-response rates for hard-to-count groups, any lessons from specific strategies used during the self-response period or in NRFU, and implications for reaching hard-to-count households in 2020. For example, what was the use of the non-ID option among different racial/ethnic groups? What were the response patterns for highly-mobile young adults and college students? How well did the Providence self-reporting rates correlate with the modeled low-response scores and the 2010 response rates?

   CENSUS RESPONSE:
   The Census Bureau accepts this recommendation. Analysis of data from the 2018 End-to-End Census Test is underway. We expect to have results available beginning later this spring. We

would be happy to share the results of our analysis, as they become available. Our analysis will include information on topics such as:

- *self-response rates;*
- *demographic information on self-respondents;*
- *pattern of demographics by mode (Internet, paper, phone); and*
- *lessons learned.*

Our analysis of hard-to-count groups will be limited to the demographics we obtain from the response data.

2. The committee would like to hear an update on the technical aspects of the communications strategies at the 2020 census at the Spring meeting, with a focus on how the data-driven strategies alluded to in the 2020 Census Integrated Communications plan will improve response rates on hard-to-count populations.

   CENSUS RESPONSE:
   The Census Bureau accepts this recommendation. We will be happy to discuss how the communications and partnership program will use data to reach hard-to-count populations and motivate response.

3. Given the stated importance of the Partnership program, we would also like to hear the progress of the Partnership program at the national and local level and how the use of systems such as the customer relationship management and the customer engagement management will improve the effectiveness of partnership or media outreach to the hard-to-count groups.

   CENSUS RESPONSE:
   The Census Bureau accepts this recommendation. We will be pleased to discuss the integration of key technology systems like the CRM and CEM into the development of our partnership program and how we focus on hard-to-count groups.

4. For both the communications and the partnership programs, we would like more information about the leading indicators through 2019 and early 2020 in these areas that would signal the need to spend more resources in one area or another.

   CENSUS RESPONSE:
   The Census Bureau accepts this recommendation. As part of our conversation at the Spring meeting, we will be pleased to discuss the data points and indicators used to allocate resources – including allocation of staff time to reach areas where the evidence shows lagging response.

5. The Census Bureau should determine the hard-to-count groups or areas that they will actively monitor during the Census collection period (considering historic patterns, CBAMS results, end-to-end test results, stakeholder input, etc.). We would be interested in learning the methods and processes that will be used to monitor progress during data collection and the planned range of responses when a lagging self-response rate is identified for a specific population.

   CENSUS RESPONSE:
   The Census Bureau accepts this recommendation. We are in the process of determining and/or finalizing plans associated with monitoring self-response rates and our response options when a lagging self-response rate is identified. These plans will be integrated in our campaign optimization strategy in which we will use components from the Integrated Partnerships and Communications Program to address issues with HTC populations. Targeted efforts - such as additional paid advertising, social media messages, or partnership activities - will be used to address issues. We will take into account which targeted efforts have worked best for the particular audience. We would be happy to share this information with you at a future meeting.

6. The committee recommends that partnership efforts should particularly target national organizations that represent groups not covered in the CBAMS focus groups or other related research, like parents of young children or others.

   CENSUS RESPONSE:
   CBAMS was a robust research study that demonstrated the motivators that work best and the barriers to response that exist for the entire population. We also have extensively tested our platform and tagline to ensure resonance across languages and cultures. Similarly, we will test our creative materials. The subjects of those focus groups were incredibly diverse. Some of our research focus groups include parents of young children. Our partnership efforts are focused on those organizations that can help the Census Bureau gain access to hard-to-count populations and that are trusted voices to those groups.

7. There are many government and quasi-government entities trying to reach the Hard to Count population, some at roughly the same time of the year. The Census should take the lead in putting together an interagency task force on outreach strategies, so that all such entities can learn from each other, at a minimum, and form partnerships when synergies exist. Some of these entities include state offices in charge of SNAP; HHS offices in charge of Healthcare.gov; the IRS through its EITC outreach; and voter turnout organizations operating within the various states.

   CENSUS RESPONSE:
   The Census Bureau is working actively across federal and state government agencies to streamline efforts and work together to communicate to jointly-served populations about the

2020 Census. The creation of an interagency task force at this stage is not feasible; however, we are working with state offices like Supplemental Nutrition Assistance Program and other federal agencies to exchange best practices and build efficiencies.

8. Promotion and guidelines for using ROAM to target outreach efforts should encourage local stakeholders to review the patterns in low-response scores in light of local knowledge about changing neighborhoods (which may not be captured in the 5-year ACS).

   CENSUS RESPONSE:
   The Census Bureau accepts this recommendation. In fact, this recommendation is one of the ideas that we keep in mind when considering how our local partners can help the 2020 Census.

9. CSAC recommends that the Census Bureau explore strategies to reach the mobile, digitally-native, young adult population who may not pay attention to physical mail. Ideas include collaborating with social media influencers or partnerships with major national banks (since they are likely to be in online banking).

   CENSUS RESPONSE:
   The Census Bureau accepts this recommendation. Our partnership and communications program is already focusing on this population and pursuing many of these suggestions as a way to reach them.

10. Protocols for item non-response should take into account the risk of affecting the willingness of other individuals to respond. For example, if a social media story were to circulate about an enumerator visit following up on a missing response to the citizenship question that could deter others from answering at all.

    CENSUS RESPONSE:
    The Census Bureau thanks the CSAC for this recommendation. A complete and accurate census is dependent upon the collection of quality data for every household in our self-response areas. We will encourage all respondents to respond to all questions on the 2020 Census. However, we recognize that with any survey or census, item non-response is a reality. While we strongly encourage respondents to answer every question for every person (for the online questionnaire, we include prompts if the respondent has failed to answer a question), we will still allow the questionnaire to be submitted even if all questions do not have a response. Failure to answer questions on the census increases the likelihood of a follow-up visit by a census enumerator. After the total person counts have been established, missing item data are imputed to ensure that all persons have characteristic values for the purpose of tabulating other census information products, such as the PL 94-171 redistricting data.

11. The answer choices to the citizenship question separate U.S. citizens born in the 50 states from other U.S. citizens born in the territories. In addition, the population of undocumented immigrants in the territories is likely to be different from undocumented immigrants on the

mainland, in terms of a wide array of variables that might make a difference in terms of NRFU impact, such as basic demographic variables, but also in terms of country-of-origin, method of entering the United States, amongst others. As a result, the NRFU impact of the citizenship question may be different in the territories. In light of this situation, the Census Bureau should consider choosing one of the territories for an ACS-based citizenship test that mimics what it is planning to do in the rest of the Nation.

CENSUS RESPONSE:
The Census Bureau thanks CSAC for this recommendation and would like to share information that may clarify how we will conduct the 2020 Census in the Island Areas/territories. When we refer to the Island Areas we include: American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, and the U.S. Virgin Islands. The Island Areas Censuses differs from the stateside Census in a number of ways including, but not limited to:

- *Inclusion of a citizenship question similar to previous censuses of the Island Areas, unlike stateside.*
- *Use of the American Community Survey (ACS) questionnaire with minor wording changes to accommodate time reference differences, incorporation of the final 2020 Census questionnaires taking into account Island Area local government concerns, where possible.*
- *Enumeration resulting from field enumerators first listing addresses using paper address registers and paper maps and for every living quarter the enumerators visit, conducting an interview with household members and following up as necessary, similar to the stateside Update Enumerate operation.*

As a result of our ongoing collection of citizenship for the Island Areas and the planned enumeration methodology for the 2020 Census, the Census Bureau does not accept this recommendation.

12. Importantly, methods to assess the quality of the 2020 Census after it is completed should be fully defined well *before* the counting begins. These may include a post-enumeration survey and comparison with other population estimates. It is possible that stakeholders will claim the Census 2020 results are not sufficiently accurate. The Census can potentially ward off these challenges by stating in advance their definition of an accurate Census 2020.

CENSUS RESPONSE:
The Coverage Measurement (CM) Program activities planned for the 2020 Census will provide estimates of net coverage error and components of census coverage for housing units (HUs) and people in HUs. The CM Program goals will be met by conducting a post-enumeration survey, designed to (1) Provide measures of net coverage error; (2) Produce measures of components of census coverage, including correct enumerations, erroneous enumerations, imputations, and omissions, and (3) Produce these measures of coverage for demographic groups and key census operations. To achieve these goals, the 2020 CM Program will design and conduct the 2020 Post-

Enumeration Survey (PES). Similar to the 2010 Census Coverage Measurement (CCM) Program and the 2000 Accuracy and Coverage Evaluation (A.C.E.) Program, group quarters facilities and people residing in those facilities are not within the scope of the 2020 CM Program.

The 2020 PES estimation process will continue to use the dual system methodology for net coverage error estimation. This methodology was used in the coverage measurement programs for the 1980, 1990, 2000, and the 2010 censuses (Fay et al. 1988, Hogan 1993, U.S. Census Bureau 2004). A detailed operational plan for the 2020 PES will be published in 2019. In addition to the PES, the Census Bureau also depends on its Demographic Analysis (DA) program to provide an important quality evaluation of the decennial census. The DA program develops national-level estimates of the population for comparison to the decennial census counts. DA estimates are developed independently from the census historical vital statistics, estimates of international migration, and other data sources. The DA estimates are then compared to the census and differences are examined by age, sex, and limited race groups. DA is the most cost effective evaluation of the census and is the only evaluation that the Census Bureau is able to release before the actual census results.

V. Use of Census Data for Disasters

The Census Bureau has done a great job at innovating in the area of data to support disaster preparedness and response. The repurposing of On the Map and Census Business Builder have gone a long way to inform stakeholders involved in disaster response. To enhance these current products, CSAC recommends the following:

1. Make sure that OnTheMap for Emergency Management includes smaller scale disasters, including tornadoes. Consider the inclusion of state-mandated disaster declarations.

   CENSUS RESPONSE:
   Smaller scale disasters, such as tornadoes, create geographic scale problems for OnTheMap for Emergency Management. The smallest geography for statistics published by the Census Bureau (and thus included in OnTheMap for Emergency Management) is the census block. The scale of some events (e.g. tornadoes, which can destroy one house but leave the next one intact) is almost always smaller than the scale of a census block and it would not be appropriate to perform a simple tabulation of the population statistics from the census blocks that are crossed by such an event. A brief description of this issue can be found here:
   https://www.census.gov/ces/pdf/2013_CES_Research_Report.pdf, Page 13, Text Box 2-1.

   Alternative approaches, such as special tabulations of confidential census data, are technically feasible; however, these present additional challenges, especially focused on disclosure avoidance, which would require a significant research and development effort.

Inclusion of state-mandated disaster declarations within OnTheMap for Emergency Management is technically feasible if the states provide automated information feeds on those events. General requirements for inclusion of specific types of events are given in the report mentioned above (same Text Box 2-1). Additionally, funding for development activities to include these new events in the application would need to be identified.

2. Work to include the US territories in OnTheMap for Emergency Management. Consider prioritizing areas that are not "low hanging fruit" in order to optimize equitable coverage of OnTheMap resources.

   CENSUS RESPONSE:
   Puerto Rico already is included in OnTheMap for Emergency Management, in that there are Decennial Census and ACS statistics (but not LODES) for the territory and some of the event sources cover Puerto Rico, including FEMA Disaster Declaration Areas and Hurricanes and Tropical Storms. The U.S. Virgin Islands are marginally included in OnTheMap in that some event sources cover the area of the islands, including FEMA Disaster Declaration Areas and Hurricanes and Tropical Storms. However, LODES and ACS do not have coverage in the Virgin Islands, and the Decennial Census creates a special summary file for the Virgin Islands, which is not part of the standard SF1 release (used to supply data for OnTheMap for Emergency Management). One possible action would be to investigate whether the U.S. Virgin Islands Summary File is compatible with the data currently provided from SF1, and if so, include that data in OnTheMap for Emergency Management. Funding for development activities to include these new data in the application would need to be identified.

   Other island areas, such as American Samoa, Guam, and the Northern Mariana Islands, do not have any coverage in OnTheMap for Emergency Management with the possible exceptions of FEMA Disaster Declaration Areas (although none have been observed in the historic events tracked by OnTheMap for Emergency Management). Those island areas do have special summary files related to the 2010 Decennial Census, but those data have not been integrated into the application. Specific use cases and data sources for including the island areas would need to be developed and evaluated to determine feasibility and cost.

3. Make damage layers in OnTheMap for Emergency Management downloadable so that localities and local organizations can overlay them on local data sets.

   CENSUS RESPONSE:
   Event boundaries are currently exportable as KML files. Data associated with each event are currently exportable as CSV files, in both "raw" form and in a summarized form similar to what is seen on screen in the application.

   Further download/export options would need to be proposed and evaluated to determine feasibility and cost.

Substantial gaps in disaster recovery data remain. In order to begin innovating in this area, CSAC recommends the following:

4. Establish a task force or a working group with FEMA, on-the-ground post-disaster data experts, and other disaster recovery stakeholders to better understand and catalog their data needs to support and inform disaster recovery in the months and years after a disaster.

   CENSUS RESPONSE:
   Currently, the Census Bureau's Emergency Preparedness and Response Team is a member of the following FEMA and other Federal Agency Working Groups:

   FEMA Interagency Recovery Support Function Working Groups
   1. *Economic Recovery*
   2. *Community Planning and Capacity Building Recovery*
   3. *Data Analysis Recovery*
   4. *Health and Social Services Recovery*
   5. *National Business Emergency Operations Center: Economic Assessment Team*
   6. *Modeling Data Working Group*

   FEMA Interagency Emergency Support Function Working Groups
   In 2019, the Emergency Preparedness and Response Team has begun additional outreach to FEMA to become incorporated into as many of the 15 Emergency Support Function Working Groups where Census Bureau data would be relevant to their missions.

   White House
   1. *U.S. National Science and Technology Council: Subcommittee on Disaster Reduction*
   2. *U.S. National Science and Technology Council: Subcommittee on Disaster Reduction Working Group on Flood Inundation and Mapping*

5. Research the acquisition and dissemination of alternative data sources for population trend indicators that prioritize timeliness with some level of geographic granularity (sub-state areas). There are precedents for federal agencies enabling access to proprietary data, such as the ESRI-provided consumer spending in Census Business Builder or HUD's publishing of tract-level USPS occupied and vacant addresses. For example, in the case of disasters, the Census could negotiate with the USPS to publish monthly data at the tract level or quarterly data at the block group level for the affected area, neither of which is likely to threaten the USPS sales of record-level data.

   CENSUS RESPONSE:
   The Census Bureau is committed to providing the most accurate population estimates possible. To do this, we have developed an annual cycle using administrative records, the cohort component method, a top-down approach, and a great deal of quality control.

Though we typically use Internal Revenue Service tax returns to estimate domestic migration and American Community Survey data to estimate international migration, we have supplemented this with other data in the past. We used United States Postal Service (USPS) data to adjust domestic migration in vintage 2006 estimates after Hurricane Katrina and Airline Passenger Traffic data to adjust migration in and out of Puerto Rico in vintage 2018 estimates after Hurricane Maria.

There are two major limitations with releasing monthly population estimates for sub-state areas: data quality and process. More research would be needed to further assess the quality of alternative sources of data. Our production of estimates requires processing higher aggregation data (nation before states, totals before characteristics, etc.) as controls for the variability in smaller cells. Retooling the estimates production process to focus on monthly data would have a major impact on our program. Resources are not available to further assess this proposal.

6. Consider the possibility of developing standard methodologies for population displacement surveys, to be rapidly deployed in the aftermath of a disaster.

   CENSUS RESPONSE:
   ADDP has had some experience with adding questions to a survey very quickly after a natural disaster. On August 29, 2005, Hurricane Katrina made landfall on the Gulf Coast. During the September CPS data collection, field representatives attempted to contact all sample household in the disaster area except those in mandatory evacuation areas. In October, questions were added to the CPS instrument to identify people who were evacuated from their homes and were staying in CPS households outside of the evacuation area. These types of questions can be added to the CPS again, should there be another natural disaster that displaces a large number of people.

7. Leverage the use of remote sensing technologies to produce timely analysis of disaster recovery.

   CENSUS RESPONSE:
   The Emergency Preparedness and Response Team works with FEMA and others to better understand how new and emerging technologies can be used. There are many federal agencies involved in this area including FEMA, The Army Corp of Engineers, NASA, NOAA, United States Geological Survey, and others.

   Additionally, the Census Bureau continues to explore alternate ways to measure statistics across the demographic and economic programs.

8. In order to improve the timeliness of data necessary to respond to a natural disaster, CSAC requests that Census Bureau research the potential of performing surveys using text messaging devices in the aftermath of a disaster, including any special authorizations that are needed, and report back to CSAC.

CENSUS RESPONSE:
The Census Bureau continues to research innovative ways of contacting sample cases outside of the USPS. Text messaging is among the options, but more research is needed to ensure we can accurately link phone numbers to addresses, determine the best person to contact in a unit, and investigate the policy implications related to the use of text messaging services. The 2020 Census team is considering an experiment with text messaging, but are still working through the details. We hope to learn the implications of using text messaging from the 2020 experience so that we can consider it for future efforts; not only as a method for communication after a natural disaster, but as a cost-effective means of securing response.