

THE STEERING COMMITTEES OF

CENSUS INFORMATION
CENTERS

FEDERAL STATE
COOPERATIVE FOR
POPULATION ESTIMATES

STATE DATA
CENTERS

November 27, 2019

Steven Dillingham, Director
U.S. Census Bureau
4600 Silver Hill Road, Room 8H001
Washington, DC 20233

Dear Director Dillingham,

The members of the Census Bureau's key partnership programs (Census Information Centers (CIC), Federal-State Cooperative for Population Estimates (FSCPE), and State Data Centers (SDC)) present this letter outlining our questions and concerns regarding the Census Bureau's move to a new Differential Privacy Disclosure Avoidance System (DAS). As the Bureau's premier partners and supporters, the three networks want to ensure that we fully understand the Bureau's decision process. We have concerns that this implementation has been driven by data scientists with limited consideration for users' needs. We are particularly concerned that insufficient analysis has been conducted regarding how DAS will affect the Census data used for informing policy and allocating public and private funds. We hope to broaden the discussion and raise awareness of the impacts of DAS on state and local decision-making.

Our network members understand the Bureau's objective to balance privacy with data usability and availability. However, the repercussions and loss of data should be carefully weighed and evaluated relative to the Bureau's responsibilities for privacy protection under 13 USC § 9. Our committees and networks have concerns and questions regarding the timing, breadth, and scope of DAS as implemented by the Bureau. Further, we have concerns that the proposed implementation violates the Census Bureau's obligation, under 13 USC § 141, to provide a Redistricting Data File with accurate population counts. It is only recently that we have had enough specifics to analyze the potential impacts and provide meaningful feedback to the Bureau. The data released however is not complete as many issues are not yet resolved.

Our three partnerships are fully engaged in public outreach and promotion of the 2020 Census. Our communications have emphasized the goal of reliable public data. If DAS impacts the numbers to the point that the data is not fit for use, or not available for small communities, then the Census Bureau's standing as the gold standard for data will be diminished.

We are asking that the Bureau provide a detailed response and time for discussion of the points we raise in this letter. We would like to have this discussion before year-end. There is considerable concern about Differential Privacy DAS, its origins as a policy, and its implementation. This extensive change will only be successful with full vetting across the user communities and the Bureau's other federal partners. We hope that this discussion with the Bureau can produce a policy direction and DAS implementation that serves public data users before there are adverse effects.

Thank you for your attention to our questions and comments. As the Census Bureau's key partners, major users and disseminators of census data, we look forward to your responses and discussions. Our steering committees can be reached by email: jhardcastle@tax.state.nv.us, or by US mail: Jeff Hardcastle, State Demographer, 4600 Kietzke Lane, Building L, Suite 235, Reno, NV 89502. Our network members are here to support the Census Bureau's work and the success of the 2020 Census.

Sincerely,

CIC Steering Committee
FSCPE Steering Committee
SDC Steering Committee

cc: John Abowd, Associate Director for Research and Methods
Stephen L. Buckner, Assistant Director of Communications
Misty L. Reed, Division Chief, Customer Liaison & Marketing Services Office
Lakiva N. Pullins, Assistant Division Chief, Customer Liaison and Marketing Services Office
Leland Todd Webb, Chief, Data Users Branch
Karen Battle, Chief, Population Division
Michael B. Hawes, Senior Advisor for Data Access and Privacy, Research and Methodology Division
Marc J. Perry, Senior Demographic Reviewer, Population Division
Rachel Marks, Senior Technical Expert on Population Statistics, Population Division
Kevin Barragan, Statistician/Demographer, Coordination, Dissemination, and Outreach Branch, Population Division
CIC Steering Committee and Members
FSCPE Steering Committee and Members
SDC Steering Committee and Members

Attachment: (3 pages)

Questions Regarding the Proposed Disclosure Avoidance System (DAS)

- a) *To date, what was the process used for input in making the decision to implement DAS?*
- b) *What inputs (testimony, research, outside experts, etc.) did the Bureau's Data Stewardship Executive Policy Committee/Disclosure Review Board use to decide the final optimal privacy-loss budget (trade-off between privacy loss budget and data accuracy)?*
- c) *Can you provide the planned scope of DAS project, the datasets and programs that will be affected by DAS (near and long-term), and the implementation schedule?*
- d) *Does the Bureau have research that shows the 2010 Census file could be reconstructed / individuals identified **without** use of outside data files?*
It would be helpful for us to know if the reconstruction could be done without the use of a commercial dataset(s); and, if not, what commercial datasets the Census Bureau has used in reconstruction exercises.
- e) *We understand that the Bureau is considering other formal privacy systems for public data products that include tables for detailed race and Hispanic origin tables, family/household tables that were included in Summary Files 1 and 2 in 2010, an American Indian and Alaska Native Summary File, and the Public Use Microdata Sample (PUMS) File. What are the other formal privacy systems under consideration for these data products? When will information about these systems and the availability of products be released and what will be the opportunity for input by the networks?*
- f) *How is the Bureau coordinating the implementation of DAS across divisions and branches? What are the contingency plans if DAS cannot be implemented as currently envisioned?*
- g) *Can you provide specific information regarding the criteria that will determine suppression by geographic level?*
We are concerned that data may not available for key geographies including but not limited to blocks, block groups, tracts, census designated places, and minor civil divisions. It is of grave concern that DAS would have a substantial adverse impact on the availability and quality of data for small communities.
- h) *Can you provide similar information for other programs (American Community Survey, Economic Census, National Center for Health Statistics, etc.) on the tables or products that will be either modified or suppressed by using DAS or similar techniques, and the impact on geographies by FIPS and by NAICS for business data?*
- i) *Will DAS also be implemented on data products produced for other agencies? Have these other agencies (such as BEA, the Department of Housing and Urban Development, the Bureau of Labor Statistics, the National Center for Education Statistics, and the National Science Foundation) and their data users provided comments? If so, in what form and where can the comments be accessed?*
- j) *Will DAS be implemented on data products produced as a special tabulation? Has there been any research by the Bureau on special tabulation uses and the impact DAS has on these tabulations?*
- k) *What criteria were used to decide on the proposed suite of the demographic and housing characteristic file (DHC) data tables?*

- l) Has the Census Bureau considered expanding the privacy budget as the 2020 Census data becomes less current and therefore less valuable to potential reconstructors? Typically the Census Bureau does not issue Summary File 2 data until 2 or more years after the census is taken. Is there consideration in balancing privacy, currency, and accuracy in the implementation of the DAS?*
- m) For the tables that will be released for which DAS has been applied, what impact will it have on the ability to compare tables over time?*
The networks have already identified unacceptable inconsistencies when comparing the demonstration products with previously published 2010 data.
- n) Has the Bureau researched the impact of not producing data that was previously available? Has the Bureau considered that federal, state, and local programs in many cases, have statutory or administrative mandates requiring the use of Census data for funding and for reporting?*
Many of these requirements have been developed because accurate Census data provides a way to equitably distribute state and local revenues and services.
- o) Most Census survey data already have margins of error (MOE). Does applying DAS compound these errors?*
The three networks already hear many complaints about MOEs – particularly for small geographies. We are concerned that, not only will the data be less usable, but that Census survey respondents will be unwilling to fill out forms if they perceive the resulting data is less accurate and unhelpful for their needs.
- p) Why did the Bureau take the proactive stance to be the global leader in implementing disclosure avoidance, when so many vendors are collecting, selling, and publishing data that are not under the Bureau's control?*
It is important that our network members, data users, and the stakeholders we serve understand why the Bureau took the action proactively to be the global leader in disclosure avoidance without, as far as we know, any major challenge that the Bureau was not upholding its 13 USC mandate. It is also important to acknowledge that the Bureau's initiative will not solve the global problem of personal data disclosure.
- q) What other methods or consequences did the Bureau consider for protecting privacy, either legal or methodological, which would fulfill the Bureau's duty to protect an individual's record(s) and still produce data that can be used by the everyday user and local elected officials? How is this being handled by other governmental statistical agencies both within and outside of the US?*
- r) Has the Bureau considered the consequences of the implementation of DAS on non-governmental entities and programs that provide key community services? Has the Bureau presented the impacts and received input from small states, local, and non-profits on the implementation of differential privacy?*
A major concern is whether these organizations will have usable data to conduct research, present a case for grant funding, and build the right sized and type programs needed for their communities.
- s) Has the Bureau determined the impact on program reports whose findings may be distorted due to the implementation of DAS and may no longer accurately represent the reported geographic area, population group, or economic sector?*
- t) Has the Bureau considered that the implementation of DAS will result in limited data availability for small geographies, leading these entities or service providers to purchase data or conduct surveys through private companies?*
The profusion of companies willing to provide data and surveys may compound disclosure issues since they are not subject to 13 USC requirements and will not use the same strict methods and guidelines the Bureau employs for both data collection and tabulation. For example, with the lack of state level population

projections the private sector has stepped in from a variety of vendors with different products and levels of transparency.

- u) *What plan does the Bureau have to inform and work with data users to ensure the implementation and impacts of DAS do not have negative consequences?*
- v) *After data release, how is the Census Bureau going to handle criticism from the public that starts questioning the quality of the data because they find implausible numbers or don't recognize themselves or the area they live in in the published numbers? Will our networks get any guidance on how to deal with that criticism?*
- w) *What is the process, format, and timeline for the three networks to provide input to the Bureau?*
The input would include both the specific impact of DAS on the data for governmental and non-governmental organizations, as well as the result of network member comparative analyses of the demonstration tables and the 2010 tables.
- x) *The Census Bureau has asked each of the partnerships to provide support for the implementation of differential privacy. Can you please provide what you are requesting each partnership to do to show support?*
- y) *How does injecting noise into the data, disconnecting household relationship and effectively changing population counts for small areas impact the Census Bureau's residence rules and how local and state governments review the accuracy of the Census?*
- z) *How does the Bureau justify shrinking the availability of data about communities ranging from Asian ethnic groups and the Middle Eastern community when they have been asking for an expansion of how their specific communities are reported at least at a national level?*