

**UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF NEW YORK**

Dated: August 9, 2018

Respectfully submitted,

CHAD A. READLER
Acting Assistant Attorney General

BRETT A. SHUMATE
Deputy Assistant Attorney General

JOHN R. GRIFFITHS
Director, Federal Programs Branch

CARLOTTA P. WELLS
Assistant Branch Director

/s/ Martin M. Tomlinson

KATE BAILEY
GARRETT COYLE
STEPHEN EHRLICH
CAROL FEDERIGHI
DANIEL HALAINEN
REBECCA KOPPLIN
MARTIN M. TOMLINSON
Trial Attorneys
United States Department of Justice
Civil Division, Federal Programs Branch
20 Massachusetts Ave., N.W.
Washington, DC 20530
Tel.: (202) 353-4556
Fax: (202) 616-8470
Email: martin.m.tomlinson@usdoj.gov

Counsel for Defendants

**UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF NEW YORK**

STATE OF NEW YORK, et al,

Plaintiffs,

v.

UNITED STATES DEPARTMENT
OF COMMERCE, et al,

Defendants.

No. 1:18-CV-2921 (JMF)

CORRECTED DECLARATION

JOHN MARON ABOWD, Ph.D., makes the following Declaration under the penalty of perjury pursuant to 28 U.S.C. § 1746, that the following is true and correct to the best of his knowledge and belief:

1. I am the Associate Director for Research and Methodology and Chief Scientist of the United States Census Bureau.

2. Section 9 of Title 13 ("Title XIII") governs the protection of statistical data used in the development of the census and all Census Bureau data collection activities. Title XIII precludes the Secretary of Commerce, his employees, or any person who has sworn to protect this sensitive information from making "any publication whereby the *data* furnished . . . can be identified." (emphasis added). In *Baldrige v. Shapiro*, the Supreme Court interpreted the language of Section 9 literally, emphasizing that the language places confidentiality protections on the data itself. 455 U.S. 345, 356 (1982) (observing that "the language of each section refers to protection of the 'information' or 'data' compiled . . . By protecting data revealed 'on behalf of' a respondent, Congress further emphasized that the data itself was to be protected from disclosure. The legislative history also makes clear that Congress was concerned not solely with protecting the identity of individuals").

3. The language of Sections 8(b) and 9 express an explicit congressional intent to preclude all raw census data from the reach of civil discovery. *Id.* at 361.

4. Disclosure avoidance is the Census Bureau's process for protecting the confidentiality of data, as required under Title XIII.

5. Before any product/document involving census data may be released, the product/document/data to be produced or released must be reviewed by the Census Bureau to ensure that no identifiable Title XIII confidential data are or may be disclosed. Should the Census Bureau's Disclosure Review Board (DRB) determine that the product/document/data

does or reasonably could result in such disclosure, then the data product will be modified prior to approval for release.

6. When the DRB is asked to conduct a review and allow the release of summary statistics or properly anonymized micro-data records, its responsibility is to ensure that the data in the release cannot be used, alone or in combination with other data publicly available, to re-identify any individual or business included in the confidential files and any data item supplied by that entity under the authority of Title XIII. This is the serious charge that the DRB was given in 1995, as the successor to the Microdata Review Panel, which was chartered in 1981. The DRB is organized by the Data Stewardship Executive Policy Committee at the Census Bureau. Its charter is published on www.census.gov at http://www2.census.gov/foia/ds_policies/ds025.pdf. The DRB meets twice weekly and continuously revises the requirements for data release in conformance with best practices for confidentiality protection.

7. When creating the DRB, the Census Bureau explicitly recognized that improperly prepared tabular summaries and microdata could result in harmful and illegal re-identifications of individual or business data items.

8. Over the last 23 years the DRB has continuously updated the procedures required to secure permission to release summary tabulations and microdata in order to keep up with advances in mathematics, statistics and computer science.

9. A disclosure of data occurs when someone can use published statistical information to identify an individual who has provided confidential information or associate a specific data item collected under the authority of Title XIII with a particular individual respondent.

10. For data tabulations, the Census Bureau uses disclosure avoidance procedures to modify or remove the characteristics that put confidential information at risk for disclosure. Although a published table may appear to show information about a specific individual, the Census Bureau has taken steps to disguise or suppress the original data, while making sure the results are still useful. The techniques used by the Census Bureau to protect confidentiality in tabulations vary, depending on the type of data.

11. The prohibited disclosure of data subject to Title XIII protection can occur even in the absence of identifying information. Just because someone's specific name or address does not appear, it does not mean that certain data (or analysis about the difference between two data points) does not implicate an individual's confidentiality. One must understand that the cited statistics are, themselves, composed from samples, and the analysis of those samples changes frequently, especially during the process of confirming that the various data files have been properly extracted, loaded and linked.

12. Differences in results that arise from this process can easily expose single micro-data records and, thus, information subject to non-disclosure under Title XIII. Therefore, the primary reason such releases are barred from disclosure is to avoid any type of disclosure that could be brought about by deriving information from the known (or discovered) differences.

13. Title XIII requires the Census Bureau to account for such “complementary disclosure,” which is the release of data that does not appear to contain individually identifiable information (e.g., name, address, social security information, birth date, etc.), but could result in identifying individuals when that data is coupled with other information in existing Census Bureau publications and in publicly-available information. For example, while a preliminary percentage, standing alone, may not reveal personally identifiable information, when such a percentage is combined with other information, it can and has been used by sophisticated parties to derive personally identifiable information.

14. For the official publications of censuses and surveys, the DRB does not permit release of preliminary results, except where such preliminary results are a specific scheduled release, but even then certain disclosure avoidance procedures are followed. For example, while the Monthly Advanced Retail Trade Survey (MARTS) permits the release of preliminary results, the final Monthly Retail Trade Survey results constrain the disclosure avoidance procedures on the MARTS. The official tabulations themselves are subjected to a host of disclosure avoidance procedures, including primary item suppression, complementary item suppression, cell suppression, complementary cell suppression, whole table suppression, input noise infusion, record deletion, record swapping, synthetic records, synthetic values, and output noise infusion.

15. These are generally called “confidentiality edits” and they are documented in the technical reports accompanying the tabular summaries or public-use micro-data.

16. When ad hoc research reports are prepared from the confidential data, pursuant to approved projects, they are subjected to the same types of reviews. No preliminary analysis may be released, and the DRB may require any of the disclosure avoidance procedures used in official products before the final version of the results may be released.

17. The DRB recognizes that some research reports will be revised after they have been released in original form. In that case, it requires a thorough review of the differences between the original results and those proposed in the revision and it may again require any of the disclosure avoidance procedures used in official products.

18. The DRB process avoids the release of intermediate work product to the extent that it can be used in combination with the official publications and the final work product of the particular project to re-identify individual respondents and their data items.

19. This is not a hypothetical risk. The Census Bureau, along with other statistical agencies, has acknowledged the issues raised by the Dinur-Nissim (2003) “database reconstruction theorem,” also known as the Fundamental Law of Information Reconstruction, which says that overly accurate estimates of too many statistics can completely destroy privacy. Subsequent work has shown that stripping obvious identifiers such as names, addresses, and identification numbers is inadequate in the face of modern computational and informational resources. The cumulative effect of statistical releases can compromise the privacy of some individuals. The Census Bureau has dedicated significant resources to address this issue, and to balance the interests of data accuracy and the public-good mission to publish data that are suitable for their intended uses with the interests inherent in privacy loss.

20. Intermediate work product is routinely capable of being used, in conjunction with official publications and final work product, to re-identify individual respondents and their data items.

21. First, intermediate work product contains errors that could be used to “reverse engineer” personal information.

22. Second, the dynamic nature of extracting, editing and analyzing data from multiple internal confidential files requires test runs and preliminary reviews, which could again enable someone to re-identify individual respondents and their data items.

23. In the event that the DRB is asked to review results that were not intended for release, it denies the request when final results have already been released unless the input samples are identical, and the preliminary analysis does not by itself expose individual records by, for example, subtly changing definitions of sub-groups. If the final results have not yet been released, it warns the researcher that the final analysis will be constrained to use inputs compatible with the preliminary analysis so as to avoid these same risks.

24. The following Administrative Record documents contain redactions pursuant to Title XIII, in that the redacted figures comprise intermediate work product, as described above: Bates Nos. 10357, 10382, 10385, 10465, 10469, 10479, 10489, 10499, 10509, 10520, 10530, 10540, 10550, 10564, 10572, 10608, 10649, 10678, 10724, 10742, 10749, 10757, 10770, 10781, 10795, 10810, 10822, 10835, 10849, 10874, 10954, 10957, 10961, 10962, 10964, 10966, 10968, 10970, 10973, 10975, 10977, 10978, 10980, 10983, 10984, 10987, 10991, 10993, 10999, 11003, 11004, 10862, 10932, 10936, 10940, 10953, and 10996.

25. The following documents do not contain information subject to protection under Title XIII, but contain sensitive proprietary information technology information that is redacted as proprietary agency information: 10407 and 10995.

26. The following documents do not contain information subject to protection under Title XIII, and if they are not subject to any other privileges or protections may be available for release following DRB review, with necessary redactions: 10462, 10913, and 11025.



John M. Abowd, Ph.D.
Associate Director and Chief Scientist
Research and Methodology
United States Census Bureau