

IN THE UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF CALIFORNIA

CITY OF SAN JOSE, CALIFORNIA, *et al.*,

Plaintiff,

v.

DONALD J. TRUMP, *et al.*,

Defendants.

No. 5:20-cv-05167-LHK-RRC-EMC

STATE OF CALIFORNIA, *et al.*,

Plaintiff,

v.

DONALD J. TRUMP, *et al.*,

Defendants.

No. 5:20-cv-05169-LHK-RRC-EMC

Declaration of John M. Abowd, Ph.D.

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

Qualifications

1. I am the Chief Scientist and Associate Director for Research and Methodology at the United States Census Bureau. I have served in that capacity since June 2016. The following statements are based on my personal knowledge or on information supplied to me in the course of my professional responsibilities. These statements are in support of the Defendants' opposition to the Plaintiffs' Motion for Partial Summary Judgment.

2. In 1977, I received my Ph.D. in economics from the University of Chicago with specializations in econometrics and labor economics. My B.A. in economics is from the University of Notre Dame.
3. I have been a university professor since 1976. My first appointment was assistant professor of economics at Princeton University. I was also assistant and associate professor of econometrics and industrial relations at the University of Chicago Graduate School of Business. In 1987, I was appointed associate professor of industrial and labor relations with indefinite tenure at Cornell University. I am currently on unpaid leave from Cornell University to work in my current position at the Census Bureau as part of the Career Senior Executive Service.
4. I am a member and fellow of the American Statistical Association, Econometric Society, and Society of Labor Economists (president 2014). I am an elected member of the International Statistical Institute. I am also a member of the American Economic Association, International Association for Official Statistics, National Association for Business Economists, American Association for Public Opinion Research, Association for Computing Machinery, American Association for the Advancement of Science, and American Association of Wine Economists. I regularly attend and present papers at the meetings of all of these organizations.
5. I have served on the American Economic Association Committee on Economic Statistics. I have also served on the National Academy of Sciences Committee on National Statistics, the Conference on Research in Income and Wealth Executive Committee, and the Bureau of Labor Statistics Technical Advisory Board for the National Longitudinal Surveys (chair: 1999-2001).

Relevant professional experience

6. In 1998, the Census Bureau and Cornell University entered into the first of a sequence of Intergovernmental Personnel Act agreements and other contracts under which I served continuously as Distinguished Senior Research Fellow at the Census Bureau until I assumed my current position in 2016, under a new IPA contract. Since March 29, 2020, I have been in the Associate Director position at the Census Bureau as a career Senior Executive Service employee. While I was a senior research fellow, I worked with numerous senior executives. This includes Directors (Martha Riche, Kenneth Prewitt, C. Louis Kincannon, Stephen Murdoch, Robert Groves, and John Thompson), Deputy Directors (Hermann Habermann, Thomas Mesenbourg, and Nancy Potok), Chief Scientists (Roderick Little and Thomas Louis), and numerous other associate directors, assistant directors, and division chiefs. I also worked with Chief Economists John Haltiwanger, J. Bradford Jensen, Daniel Weinberg, and Lucia Foster, and researchers in all program areas.
7. I was one of three senior researchers who founded the Longitudinal Employer-Household Dynamics (LEHD) program at the Census Bureau. This program produces detailed public-use statistical data on the characteristics of workers and employers in local labor

markets using large-scale linked administrative, census and survey data from many different sources. The program is acknowledged as the Census Bureau's first 21st Century data product: built to the specifications of local labor market specialists without additional survey burden, and published using state-of-the-art confidentiality protection. In addition to very substantial financial support from the Census Bureau, this project was supported by a \$4.1 million grant from the National Science Foundation (NSF) on which I was the lead Principal Investigator.

8. From 2004 through 2009, I was the lead Principal Investigator on the \$3.3 million NSF-supported collaborative project with the Census Bureau to modernize secure access to confidential social science data. This project led to the first production implementation worldwide of differential privacy for OnTheMap—a product of the LEHD program. It also produced prototype confidential data access systems with public-use synthetic micro-data supported by direct analysis of the confidential data on validation servers. These projects were the precursors to the Census Bureau's current program to implement central differential privacy for all publications from the 2020 Census of Population and Housing, which will be the first large-scale production implementation worldwide.
9. From 2011 until I assumed my position as Chief Scientist at the Census Bureau in 2016, I was the Principal Investigator of the Cornell University node of the NSF-Census Research Network, one of eight such nodes that worked collaboratively with the Census Bureau and other federal statistical agencies to identify important theoretical and applied research projects of direct programmatic importance to the agencies. The Cornell node produced the fundamental science explaining the distinct roles of statistical policymakers and computer scientists in the design and implementation of differential privacy systems at statistical agencies.
10. I have published more than 100 scholarly books, monographs, and articles in the disciplines of economics, econometrics, statistics, computer science, and information science. I have been the principal investigator or co-principal investigator on 35 sponsored research projects. My full professional resume is attached to this report.

Scope of work

11. I have been asked to provide expert opinion responding to the expert report submitted in this case by Dr. Matthew H. Barreto.

Expert opinion

12. The most significant challenge to the quality of 2020 Census data is the COVID-19 pandemic. The effects of the pandemic and the multiple reprograms of the 2020 Census field operations make it nearly impossible to predict with any certainty whether any groups will be differentially disadvantaged in the final count. There are no natural or field experiments that speak to disruptions on this scale.
13. The Census Bureau's randomized controlled trial of a census questionnaire with and without a citizenship question, in June 2019, showed no statistically significant difference

in the self-response rates with and without a citizenship question. With a sample of 480,000 housing units, capable of detecting differences as small as 0.5 percentage points (see <https://www.census.gov/programs-surveys/decennial-census/2020-census/research-testing/testing-activities/2019-census-test/2019-census-test-report.html>), this test was large-scale and properly designed to measure the differential self-response rates using the 2020 Census contact and self-response protocols. The June 2019 Census Test did not inform the question of whether overall self-response might have been lower because of the possibility of receiving a question about citizenship (the macro environment), nor did it inform the quality of the overall census procedures, including and especially non-response follow-up (NRFU).

14. An overview of the changes to the 2020 Census field and post-processing operations because of the COVID-19 pandemic and the shortened time window for the NRFU operation are detailed in the declaration of Albert E. Fontenot, Jr., Associate Director of the Decennial Census Programs for Census Bureau.
15. Pursuant to the President's July 21, 2020 *Memorandum on Excluding Illegal Aliens From the Apportionment Base Following the 2020 Census* ("the PM"), and based on information currently available to it, the Census Bureau is in the process of determining the appropriate methodologies and finalizing, to the extent possible, how it may exclude illegal aliens in keeping with the stated purpose of the PM to use the data for apportionment. At this time, the Census Bureau does not know exactly what numbers the Secretary may report to the President, and it is therefore impossible to assess precisely the effects of the PM on apportionment. The Census Bureau is remaining consistent with best practices for a federal statistical agency.
16. With regard to Dr. Barreto's paragraph 14 (in which he concludes that the PM will reduce participation in the 2020 Census and reduce the accuracy of the 2020 census), paragraph 19 (in which he concludes that the PM erodes trust that many community-based organizations with experience serving immigrants built up over the past year), and paragraph 21 (in which he cites studies finding that Census participation drops in immigrant communities when federal immigration enforcement is perceived to be connected with the Census): As stated above in paragraph 12, the most significant challenge to the quality of 2020 Census data is the COVID-19 pandemic. The effects of the pandemic and the coinciding multiple reprograms of the 2020 Census field operations required to adapt to the circumstances of the COVID-19 pandemic make it nearly impossible to predict with any certainty whether any groups will be differentially disadvantaged in the final count. There are no natural or field experiments that speak to disruptions on this scale. Additionally, as demonstrated in paragraph 13 above, a randomized trial of the actual protocol used would be the best evidence to properly draw any conclusions.
17. With regard to Dr. Barreto's paragraph 82, in which he references 2018 survey research that he conducted in relation to the citizenship question on the 2020 decennial, and his conclusions that participation in the Census would increase after removing any fear of immigration status being exposed: The best way to accurately develop such conclusions

is to conduct a randomized trial of the actual protocol being used in the 2020 Census, as referenced in paragraph 13 above. His research failed to do this making his conclusions as set out in paragraph 82 less reliable.

18. With regard to Dr. Barreto's paragraph 84, in which he concludes that administrative records are less useful than direct responses: The lower linkage rates for administrative data from the groups that he mentions may be offset by other features of the NRFU and count imputation operations. More field visits by enumerators are still scheduled to occur. Because those enumerated via administrative record linkage are removed from the NRFU workload, more enumerator time is available to seek responses from the groups not well covered in the administrative records. In addition, after some number of visit attempts, the enumerator will only try to get a population count, which does not have any associated characteristics (like Hispanic ethnic origins). Such a population count does not present the same incentives to avoid responding nor to misrepresent the number of people in the household. As long as the NRFU reaches comparable levels of completeness as recent decennial census—more than 99% of MAFIDs accounted for—the count imputation system can use the local correlation that Dr. Barreto notes to control the effects of administrative record enumeration on hard-to-count populations, at least with respect to apportionment populations.
19. With respect to Dr. Barreto's paragraph 86, his assertions that the count imputation process is biased by non-ignorable non-response are speculative. In the presence of the pandemic, it is very difficult to predict which neighborhoods will have larger count imputation rates and which will have smaller ones.
20. With respect to Dr. Barreto's paragraph 88, he misuses Rubin's missing data definitions. MCAR means that no variables-measured or unmeasured can predict which units are missing. Ignorable missing data (the standard assumption used by statistical agencies) means that the observed responses can be reliably used to predict the unobserved ones. Since the observed responses are all the agency typically has (including data in the sampling frame), that is all the data it can use for imputation. Non-ignorable missing data means that some unobserved information on the non-respondents is required to accurately predict their missing responses. Such information comes from extra-survey sources. The quote from my previous testimony was taken out of context. As I have previously explained, accuracy has at least two dimensions. These are commonly called bias—the statistic's tendency to systematically overcount or undercount its target—and variance—the statistic's tendency to fluctuate around its target. Count imputation is a statistical measure not based on sampling that has both of these accuracy components. I was commenting on the variance of imputations not the bias. The Census Bureau's count imputation procedure is tested for unbiasedness before use.
21. I also understand that certain plaintiffs in this lawsuit have alleged that any method of implementing the PM on the part of the Census Bureau would require the use of statistical sampling in violation of federal law.
22. As the Chief Scientist and Associate Director for Research and Methodology at the United States Census Bureau, I am familiar with the statutory law regarding the use

of statistical sampling in the apportionment context. In particular, I understand that, as a general matter, statistical sampling may not be used for determining the population for apportionment purposes.

23. I can confirm that any methodology or methodologies ultimately used by the Census Bureau to implement the PM will not involve the use of statistical sampling for apportionment purposes.
24. With regard to Dr. Barreto's paragraphs 52-57, in which he discusses the results of the 2019 Census Test comparing self-response rates for questionnaires with and without a citizenship question, see paragraph 13 above for the main result, which Dr. Barreto did not provide. In addition, see paragraph 13 above for the main qualification—that the study does not address NRFU. Dr. Barreto's statements assume that the statistically significant self-response differentials found for certain groups would persist after NRFU that resolved the status of at least 99% of all MAFIDs in the 2020 Census workload. The best way to support such conclusions with statistical evidence is to conduct a randomized trial of the actual NRFU protocol being used in the 2020 Census, which neither the Census Bureau nor Dr. Barreto has done.

John M. Abowd, Ph.D.
Chief Scientist and Associate Director for Research
and Methodology
Bureau of the Census

Dated: _____