

Discussion

*Stephen E. Fienberg*¹

Habermann is a dedicated professional serving a lead statistical agency in exemplary fashion. His defense of the actions of those in his agency both present and past is vigorous and impressive. We are friends and more often than not agree on statistical matters, both methodological and professional. Unfortunately, in the present instance, I believe that his arguments about the past are not well supported by the facts and that his views about ethics are out of step with much of the profession and the expectations of modern American society. In this commentary I discuss the relevance of history, and matters of law and ethics, in that order.

Key words: Census data; confidentiality; data sharing; ethics; disclosure limitation.

1. History of the Census

I begin with the history of the census and group identification, based on ethnicity and race. This goes back to the first census and continuing up through the 1960 census race was measured in a variety of different ways but largely as an instrument for government exclusion of various groups, such as Blacks and American Indians. For a detailed history and information on groups identified in successive censuses, see especially Anderson (2002) and Anderson and Fienberg (1999). Thus to view what occurred at the U.S. Census Bureau in 1942 and the role of the Bureau's use of census data out of the context of this history extending some 150 years before the incident and 20 years after, is to look at the world through rose-colored glasses and ignore the societal impact of data collected by the Bureau over time. After all, the data collected on Negroes in the 1940 Census was explicitly used by the federal and state governments to reinforce segregated systems. Today we view segregation as immoral. Thus what occurred in 1942 was not an isolated use of government data to adversely affect racial groups in the U.S., although it came at a time when the U.S. Census Bureau revolutionized modern census taking and the methodology for sample surveys.

2. History of Confidentiality and the U.S. Census Bureau

My philosophy colleague and ethics expert at Carnegie Mellon, Peter Masden (2003), has noted that "maintaining confidentiality is recognized early in the Western tradition as a

¹ Carnegie Mellon University, Department of Statistics, Center for Automated Learning and Discovery, and CyLab, Pittsburgh, PA 15213, U.S.A Email: fienberg@stat.cmu.edu

Acknowledgments: The research reported here was supported in part by NSF grant EIA-0131884 to the National Institute of Statistical Sciences, and by Army contract DAAD19-02-1-3-0389 to CyLab, at Carnegie Mellon University.

binding duty but that this duty has no clear boundaries in its early inception,” arguing that its roots can be found in the Hippocratic Oath. In the context of U.S. official statistics, concerns about keeping responses to government inquiries confidential did not really arise until the twentieth century and were largely focused in the domain of business statistics—to protect respondents from inquiries from other government agencies. There was a slow shift towards protecting individual data. Indeed much of that history developed around resistance by the Bureau to efforts of other government agencies to gain access to individual census data. For details on this history see Anderson and Seltzer (2004).

Had things really changed dramatically by 1942? Perhaps somewhat, but the pressures from the outside were similar to those of the past, and the careful documentation of the interaction of those leading the bureau with others in government who wished to abridge or abolish the confidentiality of census data are compelling testimony to the position Habermann argues against. Indeed, in June of 1941 just after J.C. Capt was confirmed as Director of the U.S. Census Bureau, he helped to produce language for a bill repealing the confidentiality of census responses for national defense; fortunately the bill with this language never made it out of House committee. Following the Pearl Harbor attack, Capt even proposed appealing the provision by executive order. Clearly the politics within and outside the Bureau were lining up against preserving confidentiality. Habermann carefully avoids citing the historical narratives developed by Seltzer and Anderson (2000, 2001, 2003) based on publicly accessible archival information, and chooses to cite instead internal U.S. Census Bureau investigations, many of which are not available to the readers of the *Journal of Official Statistics*.

What are the facts regarding the U.S. Census Bureau’s role in the 1942 evacuation? I think that there is some agreement to the effect that there is no evidence that the Census Bureau and its staff, including Calvert L. Dedrick, directly violated Title 13, but there is also incontrovertible evidence that Dedrick used census punch cards to identify groups of Japanese Americans and assisted in the effort to round them up which led to their internment. There also appears to be archival evidence that Dedrick was aided and encouraged by senior census officials, including Capt. So we can take as a starting point agreement with Habermann that the U.S. Census Bureau did not violate Title 13, but we can also see concerted efforts to get around the confidentiality restrictions in the provision of data to aid the evacuation of the Japanese.

My Main Argument: First, there is an important distinction between legal statements regarding confidentiality on the one hand and codes of professional conduct or ethical principles and obligation on the other. The two are related, of course, since complying with the law is for most an ethical issue, but they are not synonymous as Habermann suggests. Rather, codes are supposed to be based on moral standards but not necessarily vice versa, and laws may be based on neither! Second, ethical standards and codes change over time. Thus the revisiting of incidents such as the role of the U.S. Census Bureau in the internment of Japanese Americans during World War II is not simply to make accusations or lay blame, but is more importantly a way to help ask questions about behavior today, in the context of a very different set of moral and ethical standards.

Simply put, I find the evidence compelling that by today’s standards the U.S. Census Bureau’s actions in 1941–1942 were unethical and invoking the name of Morris Hansen to suggest otherwise is mere puffery. I suggest that readers go read the Hansen (1989)

discussion to which Habermann refers. It consists of a single paragraph within a broader 3-page document in which Hansen took issue with a brief passage in Anderson's (1988) history of the U.S. Census Bureau referring to the 1941–1942 events, and with no reference to any other material. Then I suggest that readers contrast Hansen's words with the statement a little over a decade later by then Bureau director Kenneth Prewitt (2000):

The internment of Japanese Americans was a sad, shameful event in American history, for which President Clinton, on behalf of the entire federal government, has forthrightly apologized. The Census Bureau joins in that apology and acknowledges its role in the internment.

3. Statistical Ethics and Codes of Conduct

Normally those of us who write about privacy and confidentiality focus on the tradeoff between data utility and the risk of disclosure of respondent data. Since there are ethical obligations in support of data sharing and in protecting respondent confidentiality (again see Masden, 2003 and the 1999 ASA Guidelines), balancing these obligations can present an ethical dilemma that is not simply resolved by releasing data that have no possible risk of disclosure. But Habermann's paper is actually about a different ethical conflict, i.e., between obligations to respondents and the public and obligations to the rest of government. Here there is a much clearer choice and the matter is clarified in part by the change in guidelines for professional ethics brought about by the American Statistical Association in going from the 1989 version which emphasized the legalistic aspects of promises of confidentiality to their 1999 revision which placed a greater emphasis on shared values and the prevention of harm. In the extreme the revised guidelines note that "Where the end purpose of a statistical application is itself morally reprehensible, statistical professionalism ceases to have ethical worth," and we are encouraged to "Know about and adhere to appropriate rules for the protection of human subjects, including particularly vulnerable or other special populations who may be subject to special risks or who may not be fully able to protect their own interests."

4. The 2004 Controversy Regarding Arab Americans

Habermann appears to be perplexed by the public outcry in connection with the data on Arab Americans, arguing again that the Bureau did not violate its pledge of confidentiality and simply provided a tabulation from data already available in the public domain. Since one of the arguments constantly put forward in support of the Bureau's approach to confidentiality has been ensuring the trust of its respondents, why should we not be surprised by the reference to the 1941–1942 events and the potential harm associated with providing data on an ethnic subgroup to a government agency helping to prosecute the U.S. war on terrorism, especially in light of the ludicrous claimed purpose of the data request?

Of course the Bureau has a general responsibility to provide data to other government agencies that help inform policy making, but it also has an ethical responsibility to those that provide data to it, even when the data are effectively available already. Clearly the Bureau's willingness to prepare a special tabulation on Arab Americans did not trigger

a government response akin to that we see looking back to 1942, but it will hardly encourage this group and others to co-operate with the Bureau in the future. Had I been at the Bureau and in a position to respond to the original request, I would like to think that my answer would have been clear: “The Census Bureau shares much of its data as a public resource and you can prepare such a tabulation from file X available at Y. If your concern is regarding the signage in languages other than English at airports across the country we suspect that there are better sources of information to aid your decision-making.”

5. Ethical Education and the U.S. Census Bureau

I have been asked by my colleagues at the Bureau how I can reconcile my position on such matters with my advocacy of greater public access to U.S. Census Bureau data. My answer is simple. Sharing data is a matter of ethics and the U.S. Census Bureau’s data are a public good. So the need to provide greater access to Bureau data seems obvious to me. But I also see the possibility that some data can be misused, either by government officials or but others who access them. There is an ethical obligation not to aid and abet that abuse.

I applaud the U.S. Census Bureau for beginning to sensitize its staff about ethical matters. Such education is essential to the field. My only hope is that the Bureau can learn from incidents such as those discussed in Habermann’s paper, and that the statisticians at the Bureau begin to recognize the distinction between their evolving ethical obligations as statisticians and their legal obligations to respondents under Title 13. There is a growing literature on statistical ethics (see e.g., Lesser and Nordenhaug, 2004), in part as a result of requirements for research involving human subjects that includes surveys and censuses, and the standards of behavior are evolving. We would all do well to take note.

6. References

- American Statistical Association (1989). ASA Ethical Guidelines for Statistical Practice. <http://www.tcnj.edu/~asaethic/asagui.html>
- American Statistical Association (1999). Ethical Guidelines for Statistical Practice, Prepared by the Committee on Professional Ethics. Approved by the Board of Directors, August 7. [http://www.amstat.org/profession/index.cfm?fuseaction = ethicalstatistics](http://www.amstat.org/profession/index.cfm?fuseaction=ethicalstatistics)
- Anderson, M. (1988). *The American Census: A Social History*. Yale University Press, New Haven.
- Anderson, M. (2002). Counting by Race: The Antebellum Legacy. In *The New Race Question*. Joel Perlmann and Mary Waters, (eds). New York: Russell Sage Foundation, 269–287.
- Anderson, M. and Fienberg, S.E. (1999). *Who Counts? The Politics of Census-taking in Contemporary America*. Russell Sage Foundation, New York. Revised and updated 2001.
- Anderson, M. and Seltzer, W. (2004). The Challenges of Taxation, Investigation, and Regulation: Statistical Confidentiality and U.S. Federal Statistics, 1910–1965. Paper prepared for U.S. Census Bureau Symposium, Woodrow Wilson International Center for Scholars, March 4–5.

- Hansen, M.H. (1989). Discussion. Proceedings of the Section on Survey Research Methods, American Statistical Association, Washington, DC, 162. Available from World Wide Web: <<http://www.amstat.org/section/srms/proceedings>>
- Lesser, L.M. and Nordenhaug, E. (2004). Ethical Statistics and Statistical Ethics: Making an Interdisciplinary Module. *Journal of Statistics Education*, 12, 3. <http://www.amstat.org/publications/jse/v12n3/lesser.html>
- Masden, P. (2003). The Ethics of Confidentiality: The Tension between Confidentiality and the Integrity of Data Analysis in Social Science Research. Paper presented the National Science Foundation Workshop on Data Confidentiality, May 12–13.
- Prewitt, K. (2000). Additional Comments by Census Bureau Director Kenneth Prewitt on the Role of the Census Bureau and the Use of Census Tabulations in the Internment of Japanese Americans in World War II, e-mail to U.S. Bureau of Census staff, March 24.
- Seltzer, W. and Anderson, M. (2000). After Pearl Harbor: The Proper Role of Population Data Systems in Time of War. Annual Meeting of the Population Association of America, Los Angeles, CA, March 25. <http://www.uwm.edu/~margo/govstat/integrity.htm>
- Seltzer, W. and Anderson, M. (2001). The Dark Side of Numbers: The Role of Population Data Systems in Human Rights Abuses. *Social Research*, 68, 481–513.
- Seltzer, W. and Anderson, M. (2003). NCES and the Patriot Act: An Early Appraisal of Facts and Issues. Proceedings of the American Statistical Association, Section on Survey Research Methods. Alexandria, VA: American Statistical Association, 3153–3156 (CD ROM).

Received December 2005