



May 7, 2024

Sheleen Dumas,
Department PRA Clearance Officer
Office of the Under Secretary for Economic Affairs
Commerce Department
1401 Constitution Ave NW, Washington, DC 20230

Submitted via email: Thomas.J.Smith@census.gov

RE: Collection of State Administrative Records and Third-Party Data (Docket Number USBC–2024–0005) (89 FR 16720)

Dear Sheleen Dumas:

Asian Americans Advancing Justice | AAJC (Advancing Justice | AAJC) is a national non-profit, non-partisan organization founded in 1991. For over thirty years, we have served as the leading Asian American voice on civil rights issues in our nation's capital. Our mission is to advance civil and human rights for Asian Americans and to build and promote a fair and equitable society for all.

Over the decades, we have worked to eliminate the barriers that have historically resulted in undercounting and underreporting (or otherwise inaccurate counting of) Asian Americans in federal data collection and analysis efforts, particularly in the decennial census count. Our permanent census program monitors census policy and educates policy makers—including through testifying at Congressional hearings. We conduct community outreach and education on the surveys conducted by the Census Bureau, including running nationwide Asian American-focused campaigns for Census 2000, Census 2010, and Census 2020. Advancing Justice | AAJC has also served as a member of numerous advisory committees to the Census Bureau since 2000. Most recently, we served on the National Advisory Committee on Racial, Ethnic, and Other Populations, completing our second three-year term in August 2019. Additionally, Advancing Justice | AAJC currently co-chairs the Leadership Conference on Civil and Human Rights' Census Task Force and serves as a co-coordinator of the Census Counts campaign.

Advancing Justice | AAJC considers a fair and accurate census and comprehensive American Community Survey among the most significant civil rights issues facing the country today. Our wide-ranging efforts to promote civic engagement, forge strong and safe communities, and create an inclusive society are guided significantly by objective, inclusive data on America’s diverse communities and populations. We appreciate the importance of fact-based analyses and the need for disaggregated, detailed data on our community for the purposes of identifying disparate access and outcomes and devising effective solutions. To that end, we provide the following feedback in response to the Federal Register Notice seeking comments on the collection of state/local administrative records and third-party data to improve efficiency and accuracy in the Census Bureau’s data collections, and to improve measures of the population and economy. We believe the potential use of administrative records and third-party data presents opportunities and concerns, requiring careful data quality evaluations, privacy assessments, and further research.

Overview

The Census Bureau has taken many steps over the decades to improve the accuracy of the decennial enumeration, and we applaud innovation in all federal data collection activities. We believe this encourages creativity and forward-thinking, is a wise expenditure of public dollars, and takes advantage of scientific advancements in survey measurement. Nevertheless, the census continues to count some communities and population groups more accurately than others. This results in omissions of and differential undercounts for people of color, renters (a proxy for lower-income households), young children, American Indians living on reservations, and many rural and remote households, compared to non-Hispanic Whites, homeowners, and some older Americans.

Even when there may be an overcount at the national level, such as what was reported for the Asian American community in the 2020 census, variations within the Asian American community existed. Last year, Advancing Justice | AAJC and Demographic Analytics Advisors published a report, “The Quality of the Decennial Census for Asian American and Native Hawaiian and Pacific Islander Communities: An Expanded Approach”¹ that compared the postcensal population estimates with the decennial census to determine what census accuracy may have looked like in the 2020 Census for Asian American communities. We found that there were significant regional differences throughout the country in terms of potential net undercounts and overcounts. Therefore, we must evaluate any proposed use of administrative records from any source in the context of an

¹ <https://www.advancingjustice-aaic.org/publication/quality-decennial-census-asian-american-and-native-hawaiian-and-pacific-islander>

overarching goal to eliminate these persistent omissions and differential undercounts in decennial censuses.

Caution against Overreliance on Administrative Data

As we noted in our most recent report, “Assessing Availability and Quality of Administrative Records for Asian Americans and Native Hawaiians and Pacific Islanders: Introduction and Federal Database Analysis,”²

“[t]he increased usage of administrative records in the decennial census raises both opportunities and concerns. It is certainly true that the increased use of administrative records can decrease costs and improve some parts of decennial census operations. However, this is dependent upon which records are used, as well as how, when, and where these records are used. An overreliance on administrative records could lead to inequities in the census based on differential coverage in these records, likely exacerbating historic and persistent differential inaccuracies in the census for communities of color. For instance, a heavy reliance on IRS tax data, without other data sources to augment it, could miss those who are less likely to file taxes, such as low-income households who may not make enough to need to file taxes and who already tend to be harder to count during a decennial census. It is therefore crucial to understand the coverage of different populations in these records, as well as the information that is available about these populations in the same records.”

Racial disparities in administrative records have been documented in different contexts, such as in healthcare.³ Any disparities in administrative records would be carried over if used excessively and aggressively in the decennial census. This is particularly problematic for smaller population groups who are more likely to be missed and less likely to be captured in administrative records.

Additionally, due to the lack of standardization across administrative databases on how race and ethnicity data are reported and collected, a large portion of our community could potentially be missed through the use of administrative data. In many datasets, “Asian” responses are lumped into an “Other” category, making it impossible to determine which individuals would identify as Asian American. Outside of the data produced by the Census

² <https://www.advancingjustice-aajc.org/report/assessing-availability-and-quality-administrative-records-asian-americans-and-native>

³ See <https://www.advancingjustice-aajc.org/publication/advancing-justice-aajc-submits-comments-response-announcement-pacaahnpi-meeting-and> (page 2)

Bureau, detailed data on Asian American and NHPI subgroups are generally unavailable. As our report on administrative records found,

“Race data available from existing federal administrative data sources is far less prevalent than information such as name and address, especially when compared to survey and census data. There are approximately 342 million estimated form responses of the estimated seven billion (including collection forms representing an individual or an entire household useful for rostering) that contain race information. Approximately 208 million responses include disaggregated race for Asian Americans and NHPI and are solely sourced from survey data collections.”

As noted in [the table](#) showing how different federal agencies collect race information in their forms, the Census Bureau collects the most information on disaggregated data, but only six other agencies collect forms with disaggregated data out of the total 33 agencies that collect race and ethnicity data. This could potentially miss a large portion of our community in these surveys through the use of administrative data, particularly as it relates to detailed data collection.⁴

Of particular concern is the potential use of commercial sources for data, as commercial data are of greatly lower quality than governmental data with respect to our community.⁵ This means that any design that relies solely, or even mainly, on administrative records in lieu of self-response and/or in-person non-response follow-up, as well as the use of commercial data except for in the most limited of ways, is highly problematic.

General Principles to Guide Administrative Records Usage

Any efforts to use administrative records more extensively, especially as it relates to enumeration, must undergo rigorous testing and there must be significant engagement with census stakeholders about such usage. As the Bureau continues to explore using administrative records more extensively, the Bureau should be guided by the overarching principle of using administrative records sparingly. Furthermore, the Bureau should rely on this data only when it is confident in the quality of the data provided through the records, the coverage of the data across different communities—particularly those who are traditionally hard to count—and that the administrative record can provide responses to all questions asked in the decennial form as it designs for the 2030 Census.

⁴ While this may change in the future as OMB’s recently revised SPD 15 requires disaggregated data by default, federal agencies will need time to fully implement this requirement. The new standards provide a five-year time frame for implementation, which would not be timely for use for the 2030 decennial census.

⁵ See Burton Reist, U.S. Census Bureau, 2010 Census Match Study Report (Nov. 19, 2012), https://www.census.gov/content/dam/Census/library/publications/2012/dec/2010_cpex_247.pdf.

Equity

The acquisition, management, and use of administrative data must be centered in equity. When seeking to balance objectives for the accuracy, privacy, and efficacy of data collections and products, the agency should seek to balance these three objectives in a way that centers and prioritizes equity. In contemplating the collection of state/local administrative records and third-party data, the Bureau should prioritize persistently undercounted and hard-to-count populations and communities in the acquisition, linkage, and utilization of administrative data. By extension, this aims to reduce disparities in data accuracy and representation. That is, the Bureau should consider how well different data sources accurately include and reflect persistently undercounted and hard-to-count populations and communities. They should focus on pursuing and acquiring certain data sets that best account for these groups.

Consistency

When producing statistics for the entire nation and its governmental units (e.g., states, counties, cities), the consistency of data across states is vital to ensuring federal funds—that might be guided by statistics generated using administrative records—are allocated fairly and equitably. We are not confident that all states collect data of comparable quality and accuracy, even when they are collecting the same information. Different program implementation protocols and eligibility requirements for program participation further compromise the consistency of data from state to state in terms of the documentation and completeness of the records.

We are concerned about the likelihood of uneven data quality with respect to data on race and ethnicity, even when states are required to follow federal standards for the collection and reporting of this information. Furthermore, data that respondents might consider sensitive—such as their citizenship status, income, and household size (as well as race and ethnicity)—might be less reliable depending on the individual’s own circumstances and the broader political environment, both nationally and in each state. This may yield administrative records of uneven and even dubious quality.

Generally speaking, unless the Census Bureau can demonstrate a high level of consistency with respect to quality and reliability of an administrative data source, it should exercise great caution before using these governmental records in statistical products with legal, policy, and funding purposes. The Bureau should take into consideration whether a particular type of administrative data source is available for all 50 states--and are of comparable quality across the U.S. states. At the same time, there may be certain communities where targeted data sources from targeted states may work best to accurately include and reflect these communities; these cases should be assessed

individually to ensure equity and accuracy are met. For example, the Bureau should consider using SNAP, WIC, and TANF data or other data sources that provide coverage on communities that are likely missing from IRS records only from states where the data quality are good, even if the same level of quality is not available across all 50-states. This is especially the case for administrative data sources that might provide coverage on certain populations that are highly concentrated in certain states. For example, people experiencing homelessness are highly concentrated in a handful of states; the Bureau should prioritize getting quality and consistent data on this community from those specific states rather than a rigid rule of availability, consistency, and high-quality data source across all states.

Timeliness

The time period represented by state administrative records collected for purposes unrelated to the census becomes a paramount concern. The decennial census is unique in that it measures the location and characteristics of the entire resident population as of a statutorily established date: April 1st of the census year. State administrative records generally have no such benchmark date. Therefore, information gathered from program participants or benefit recipients could represent wide and inconsistent windows of time that diminish the usefulness of the information to produce high-quality data that supplements decennial census statistical products. This must be taken into consideration when assessing whether to acquire a particular data source and/or how to utilize it.

Conclusion

In addition to the cautions and concerns raised in these comments, it will be important for the Census Bureau to engage stakeholders proactively and meaningfully to build public trust, improve data quality and interpretation, and better understand concerns from the community. The criteria for appropriate acquisition, management, and use of administrative data should encompass shared values and expectations of data users and stakeholders. Thank you for the opportunity to comment on this important issue.

Sincerely,



Terry Ao Minnis

[Vice President of Census and Voting Programs, Asian Americans Advancing Justice - AAJC](mailto:tminnis@advancingjustice-aaajc.org)
tminnis@advancingjustice-aaajc.org