Understanding the Quality of the 2020 Census Is Essential

John Thompson

1United States Census Bureau, District of Columbia, United States of America

Published on: Feb 22, 2021

DOI: https://doi.org/10.1162/99608f92.9ce979be

License: Creative Commons Attribution 4.0 International License (CC-BY 4.0)
**Column Editor’s Note:** In this Effective Policy Learning article, John Thompson, former Director of the U.S. Census Bureau, explains that understanding and measuring the quality of the U.S. Census of Population and Housing, conducted every 10 years, is a complex undertaking. The 2020 Census has been particularly difficult due to the COVID-19 pandemic and unprecedented political interference into this scientific, statistical activity. He explores the reasoning behind the recommendations of the Census Quality Indicators Task Force of the American Statistical Association on both quality indicators and next steps for independently measuring the quality of the 2020 Census.

**Keywords:** 2020 Census, census quality indicators, pandemic census delays, census data fitness for use, census undercounts, measuring census data quality

Every 10 years, the U.S. Census Bureau conducts a decennial census of population and housing, a census that is mandated by the U.S. Constitution. The census that has been underway in 2020 (2020 Census) has faced an unparalleled set of challenges beyond the control of the dedicated professional staff at the Census Bureau. These include: the COVID-19 pandemic that greatly delayed the start of spring and summer field operations; major weather events, such as hurricanes, wildfires, and flooding that have displaced people in multiple U.S. locations and hampered field data collection operations, such as door-to-door interviews; political controversies that may have impaired counting immigrants; sudden externally driven alterations in schedules for data collection and for post-collection data processing and curation; impossible-to-meet statutory deadlines that were not formally extended; ongoing politicization around scientific activities; and litigation about the schedule and who is to be counted from U.S. state, local, and tribal governments as well as other organized groups with a stake in the outcome of the census.¹

With respect to the concerns described above, it is important to note that no census in the United States has ever been perfect. All censuses have been subject to various types of nonsampling error including misreporting, nonresponse, and recall bias. Of particular concern are undercounts, which have been measured to be larger for the Hispanic and non-White populations than for the non-Hispanic White population. For example, in the 1990 Census, undercounts of about 5% were measured for the Hispanic and Black populations compared to 1.6% for the non-Hispanic White population. For the 2010 Census, the comparable figures were about 1.5% for the Hispanic population, 2% for the Black population, and an overcount of about 0.8% for the non-Hispanic White population (Cantwell, 2012). Undercounts result from a number of factors such as high mobility, fear or distrust of government, poverty, homelessness, and nontraditional housing (Chapin, 2018). In addition, the reductions in undercounts for the 2010 Census relative to the 1990 Census were most likely the result of extensive paid advertising and local partnerships to have local trusted voices explain that the census was important for communities and that it was completely safe to respond.
Some stakeholders (e.g., the Mexican American Legal Defense and Education Fund [MALDEF], the National Association of Latino Elected Officials [NALEO], and the Leadership Conference on Civil and Human Rights) have expressed serious concerns that the unprecedented challenges that the 2020 Census has faced has affected its quality and accuracy to the point where it may not be fit for use for purposes such as the constitutionally mandated apportionment of congressional seats among the states and subsequent use by states to redraw the boundaries of new congressional districts (redistricting). There is a critical need for information to answer the many questions that have been raised around fitness for use. The American Statistical Association (ASA), recognizing these uncertainties and concerns, brought together a Census Quality Indicators Task Force (task force). The task force, made up of outside census experts, was charged with producing a set of scientifically sound, publicly available statistical indicators by which the quality (e.g., how well is race reported?), accuracy (e.g., what is the level of incorrect enumerations?), and coverage (e.g., are there large undercounts of particular population groups?) of the 2020 Census can be assessed. The indicators are intended to be used by the Census Bureau and qualified outside researchers to assess the quality of the census and share that assessment with the public.

I was honored to serve as a member of the ASA 2020 Census Quality Indicators Task Force. The following discussion will provide background on the importance and conduct of the 2020 Census, describe why the currently available information does not provide adequate quality assessments, and briefly describe the recommendations of the task force and how they could be applied to assess the quality of the 2020 Census.

The decennial census is one of the cornerstones of American democracy and has been conducted every 10 years since 1790. The U.S. Constitution requires that the decennial census be used for reapportioning among the states the seats of the House of Representatives in Congress of the United States, as well as the Electoral College. The 2020 Census will also be used for numerous other functions to support good policymaking and economic growth, including: redrawing congressional and local voting districts; allocating over $1.5 trillion of federal funds annually; informing sound policy development; providing critical information for state, local, and tribal government planning; and supplying critical information to large and small businesses to generate growth and job creation. Inaccuracies or errors in the 2020 Census such as large undercounts that are disparate for racial and ethnic groups or for particular areas will introduce serious inequities and disparities for the use of 2020 Census results for the subsequent 10-year period.

Many stakeholders depend on the 2020 Census to produce highly accurate results. In addition to the U.S. Congress, myriad stakeholders use census data for a wide array of purposes. For example: the Department of Justice and advocacy groups such as MALDEF use census data to challenge discriminatory actions in redistricting; the Department of Transportation allocates over $45 billion dollars annually to support highway planning and construction, the Department of Housing and Urban Development allocates over $16 billion in Section 8 housing vouchers; over $360 billion dollars are allocated annually to support Medicare; and the Department of Education supplies over $12 billion in education grants annually. As in many other countries,
tribal, city, and local governments use census data for planning functions that include provision of services, school location, and infrastructure decisions. Researchers and businesses also use census data for a wide array of purposes such as making informed decisions and producing representative findings on health and economic issues. For example, NORC at the University of Chicago uses census data to make sampling frames representative, and the large retail corporation Target bases decisions on where to locate stores on data from the census.

The Census Bureau is required by law to conduct an enumeration of all residents of the United States as of April 1 in the census year. The law also requires that the Census Bureau provide the 50 state numbers that will be used to reapportion the House of Representatives to the President by December 31 of the census year; and to provide the detailed data to support redrawing congressional and all other voting districts to each state by March 31 of the year following the census (for the 2020 Census these dates are December 31, 2020, and March 31, 2021, respectively).

At a very high level, the 2020 Census consists of four phases: (1) updating the list that the Census Bureau maintains of all addresses in the United States; (2) inviting each household on the address list to self-respond via internet, mail, or telephone; (3) conducting a Nonresponse Follow-up operation (NRFU) to collect information from those addresses that do not self-respond; and (4) post-collection data processing and preparation of the 2020 Census data products, including apportionment counts and redistricting data files.

Of particular concern to those outside groups depending on the census data to be as accurate as possible were actions taken by the Census Bureau to address the COVID-19 pandemic. First, the career staff at the Census Bureau prudently delayed the NRFU field operation from the planned start in mid-May 2020 to early August 2020, with a scheduled completion at the end of October 2020. To accommodate these delays, the Census Bureau requested through its parent agency, the U.S. Department of Commerce, that the statutory deadlines for apportionment and redistricting described here be extended by 120 days to allow sufficient time for the post-collection processing. However, on August 3, 2020, the Department of Commerce directed the Census Bureau to truncate the schedule to end NRFU by the end of September 2020 and to cut the post-collection data processing time by almost 50% to deliver apportionment and redistricting data by the statutory deadlines. It is important to note that in the 1990, 2000, and 2010 Censuses, the Department of Commerce had not inserted its opinions into the operational management of the decennial census. At the time of this writing, the Census Bureau has encountered a number of data anomalies in the post-collection data processing, the resolving of which has delayed the production of apportionment data well beyond the December 31, 2020, deadline, until April 30, 2021, at the earliest. While no information has been made available regarding the nature of the anomalies thus far, it is anticipated that discussions will soon be forthcoming on the anomalies under the direction of the new Biden administration.

The actions by the political leadership at the Department of Commerce to shorten the data collection and post-collection data processing schedules against the recommendation of the professional scientists and statisticians
at the Census Bureau have raised significant concerns among stakeholders regarding the quality of the 2020 Census. These concerns include whether the data are fit for use for apportionment, redistricting, and other uses described above. In addition, there were numerous stakeholder requests, including former Census Bureau Director Dr. Kenneth Prewitt (Counting Every Person, 2020) and the American Statistical Association (2020 Census Quality Indicators Task Force, 2020), for the Census Bureau to be transparent regarding meaningful, detailed measures that would describe the accuracy of the 2020 Census.

The ASA issued the report of the task force on October 13, 2020 (2020 Census Quality Indicators Task Force, 2020). The report contained recommendations for a number of indicators that could be used to assess the quality of NRFU as well as the post-collection data processing. The recommendations included producing the indicators at the census tract level to allow comparisons, where possible, with the 2010 Census to assess quality for local communities and for areas with concentrations of the population that are traditionally undercounted, such as minorities. Importantly, the task force recommended that the quality indicators, including those for census tracts, be made available to outside, independent, highly qualified researchers prior to the release of apportionment and redistricting data so that fitness for use of these data could be determined and made available to the public.

The Task Force considered three of the four phases of the 2020 Census described above and made detailed recommendations for quality indicators. A brief summary of these indicators follows:

- **Self-Response**—The Census Bureau has made extensive data available on self-response rates for states, cities, counties, and census tracts.\(^5\)
- **Nonresponse follow-up**—The task force recommended that a number of quality indicators be produced for states, cities, counties, and census tracts that could be used to assess the accuracy of the NRFU in 2020 as compared to 2010. Such indicators included the rate of households enumerated by a proxy (non-household respondent) and the rate of households for which minimal information, such as only the number of occupants, could be obtained.
- **Post-collection data processing**—The task force also recommended quality indicators be produced for states, cities, counties, and census tracts, including: the rate of duplicate enumerations and the rate of enumerations that lacked a complete first and last name.

The task force also illustrated the importance of conducting analyses of quality indicators at the census tract level. While the quality indicators for a single census tract will not be informative, census tracts can easily be associated with areas such as: those that have hard-to-enumerate populations, inner-cities, suburbs, rural, American Indian Reservations, and Native Alaskan Villages. Analysis of census tract-level data for these areas will provide valuable insights on the quality of the 2020 Census. For example, considering the two recommended NRFU quality indicators described, if large differences are detected between the 2020 and 2010 Census for these quality indicators for census tracts in hard-to-enumerate areas, then it will be likely that NRFU in the 2020 Census was not as complete as in the 2010 Census. It would therefore be likely that
undercounts will be larger in 2020 than in 2010 for these areas. In particular, if such discrepancies in NRFU between the 2020 and 2010 Censuses are disproportionately distributed for cities versus suburbs or for some states versus other states, then this could indicate that undercounts in the 2020 Census are likely to have an adverse effect on apportionment and redistricting. On the other hand, if analyses show no substantive difference between NRFU for the 2020 and 2010 Censuses for the substate areas described, then one could feel much more confident in the 2020 Census.

The Census Bureau has recognized the importance of providing information that can be used to assess the quality of the 2020 Census to stakeholders. Currently the ASA Task Force has been provided access to produce and analyze the recommended quality indicators. For the first time, the Census Bureau will make such evaluations of quality available with the release of the apportionment counts. Given the importance of the 2020 Census to our democracy and the serious concerns that have been raised regarding the 2020 Census, I am pleased to see that the Census Bureau is taking these actions and continuing the longstanding practices of openness and transparency. The 2020 Census stakeholders and the American public deserve no less.

__John Thompson is the former Director of the U.S. Census Bureau, where he served from August 2013 through June 2017. Prior to becoming Director, he spent 27 years in various positions at the Census Bureau, including serving as the senior executive in charge of the 2000 Census. Thompson also served as President and CEO of NORC at the University of Chicago. He earned a BA in mathematics and MS in statistics at Virginia Tech University.__

__Disclosure Statement__

John Thompson has no financial or non-financial disclosures to share for this article.

__References__


*Counting every person: Safeguarding the 2020 Census against the Trump administration’s unconstitutional attacks, House oversight and reform committee hearing.* (2020) (testimony of Kenneth Prewitt).


©2021 John Thompson. This article is licensed under a Creative Commons Attribution (CC BY 4.0) *International license*, except where otherwise indicated with respect to particular material included in the article.

**Footnotes**

1. The 2020 Census is not the first to experience litigation, however, it is the first for which the schedule was successfully challenged. [↩](#)

2. The House of Representatives responded by passing two separate bills to extend the legal deadlines, however, there was no such action in the Senate. [↩](#)

3. The Census Bureau announced the truncation on August 3, 2020, however, on September 18, 2020, the Officer of the Inspector General at the Department of Commerce issued a report (Gustafson, 2020) noting that the Census Bureau did not make the decision, rather, it came from the Department of Commerce. [↩](#)

4. The 2020 Census includes much more automation and use of mobile technology than the 2010 Census. Therefore, much more information including paradata, such as the time an enumerator spent to enter responses, is now available for analysis of the 2020 Census. [↩](#)

5. Census tracts are subcounty statistical areas that often represent neighborhoods. The average size of a census tract is a little below 2,000 housing units. [↩](#)