

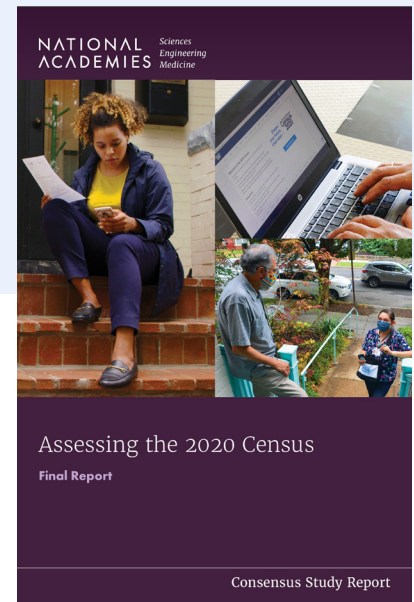
# Assessing the 2020 Census: Final Report

## THE 2020 CENSUS WAS SUCCESSFULLY EXECUTED IN THE FACE OF UNPRECEDENTED CHALLENGES, ALTHOUGH QUALITY SUFFERED AND DATA PRODUCTS WERE DELAYED

Since 1790, the U.S. census has been a recurring, essential civic ceremony in which everyone counts; it reaffirms a commitment to equality among all, as political representation is explicitly tied to population counts. Under the Constitution and the law, the U.S. Census Bureau is required to provide decennial census data to the states for redistricting and to serve as a vehicle for fairness in other arenas, such as the distribution of funds for federal programs.

Within this context, the Census Bureau charged the National Academies of Sciences, Engineering, and Medicine with convening a committee of experts to independently assess the quality of the 2020 Census and its constituent operations, drawing appropriate comparisons with prior censuses. The committee's report, *Assessing the 2020 Census*, looks at the extraordinary challenges the Census Bureau faced in conducting the census and provides guidance as it plans for the 2030 Census. The report encourages research and development as the goals and designs for the 2030 Census are developed, urging the Census Bureau to establish a true partnership with census data users and government partners at the state, local, tribal, and federal levels.

The overriding, signature achievement of the 2020 Census is that there was a 2020 Census at all. The Census Bureau and staff successfully conducted a census under exceptionally difficult circumstances—not least of which was the weeks-long delay in starting major field operations due to the COVID-19 pandemic (delayed timing that also forced census workers to contend with peak hurricane and wildfire seasons).



The report finds that the Census Bureau was able to adapt to the difficult circumstances as well as could be expected, in large part because it honed its development efforts for 2020 on a small number of innovation areas that provided the capacity for success. The report details these and other innovation areas, as well as the whole array of 2020 Census operations, and examines longstanding indicators of census data quality.

## **DATA QUALITY**

### **More Pronounced Age Heaping in 2020**

Age heaping refers to a phenomenon in which excessive numbers of people's ages are misreported as round-sounding figures, commonly those ending in 0 or 5. The report notes that age heaping is a well-recognized indicator of data quality in censuses and determines that age heaping was much more pronounced in 2020 than in 2010. The metrics provided in the report suggest that much of the age heaping in 2020 was attributable to data collected in the Nonresponse Followup (NRFU) operation, especially responses from proxies, such as a neighbor or landlord. These findings about age heaping raise concerns about the overall quality of data collected during 2020 NRFU, which the COVID-19 pandemic, among other factors, undoubtedly affected.

### **More Coverage Errors in 2020**

Every census since 1790 has missed some people who should have been counted, counted some people twice, and included others who should not have been counted. Differential under- or overcounts between particular population groups or geographic areas are more concerning than national, aggregate totals because of the implications for equity. The report examines available information from the two fundamental methods of assessing census coverage — Demographic Analysis (DA) totals and estimates derived from the independent Post-Enumeration Survey (PES) — and finds that differential coverage in 2020 affected the same population groups as in the 2010 Census but with greater magnitude. White non-Hispanic people and Asian people were overcounted in 2020 (more so than in 2010) and Black people, American Indian and Native American people, and Hispanic people were undercounted in 2020 (more so than in 2010, particularly for Hispanic people and

particularly for people who rent rather than own their residences). The report also notes that both coverage measurement techniques faced challenges in the 2020 Census — notably, the delays in census field operations delaying PES interviewing even more — and suggests research and analysis recommendations in both.

## **CENSUS OPERATIONS**

### **Master Address File**

Over its entire existence, the fundamental concept of the U.S. census has been to count each person in the right place, at their location of usual residence. Accordingly, developing an address list is a critical component of the quality of the modern census. Previous U.S. censuses constructed address lists anew each decade, but the 2000 Census began the practice of maintaining an ongoing Master Address File (MAF) to support the Census Bureau's household survey programs as well as the decennial count. The MAF is regularly updated based on information from the U.S. Postal Service's Delivery Sequence File, among other sources, and address and geographic data provided by state, local, and tribal partners.

The report draws from unprecedented access to data on the origin of MAF entries, from those on the first MAF to support the 2000 Census and those added by 2020 Census operations, and on the disposition of addresses in the 2020 Census. Upon analysis, the report concludes that the Census Bureau's routines for filtering and updating the MAF are sound as long as efforts are maintained to review and improve their efficacy. The report also recommends that the U.S. Census Bureau continue improving ways to work with user-supplied input/data resources and make it easier for state, local, and tribal authorities to supply input.

### **Nonresponse Followup**

In the 2020 Census, there was substantial geographic variation in the proportion of households that self-responded and, accordingly, in the proportion enumerated by NRFU. The report concludes that NRFU enumerations produce more item nonresponse and poorer coverage than self-responses. This result is most relevant for redistricting, where state and local officials are left to use data that vary in how accurately they reflect the age, racial, and ethnic composition of the population. The report recommends that the U.S. Census Bureau consider a major reduction in the use of proxy interviewing for NRFU, if not the elimination of proxy reporting in all but very limited circumstances, when alternative information of sufficient quality is available.

### **Use of Administrative Records for Enumeration in the 2020 Census**

Administrative records data have been used in previous censuses for many purposes, including as input to updating the MAF, in coverage evaluation, and as a method to count military personnel overseas. However, the 2020 Census was the first to use administrative records data to enumerate nonresponding households in limited circumstances — when at least one enumerator visit also failed to produce a return and when the records information was deemed sufficiently reliable. The report concludes that this use of administrative records to supplement census operations was a successful innovation and recommends further research and refinement. However, it strongly cautions that major movement toward an administrative records-based census — enumerating a substantial proportion of the population without a contact attempt — is a challenging and long-term proposition that is not feasible for the 2030 Census.

### **Enumeration of Group Quarters**

Measuring the population in group quarters (GQ), where people live or stay in a group-living arrangement that is owned or managed by organizations providing housing or services for residents, is problematic in every census — but vastly more so in 2020, with the COVID-19 pandemic. As with other 2020 Census operations, Census Bureau staff exerted laudable and extraordinary efforts to adapt to the conditions and make the GQ count as complete as possible, with the job made more difficult by the

dispersal of the college/university student population and the lockdown of nursing and health care facilities due to the pandemic. The count of students living in college/university student housing was complicated by the mismatch between census content and the U.S. Department of Education’s interpretation of permissible data release under federal privacy law (the Census Bureau’s efforts to steer schools to report electronically clashed with the guidance that schools could only report limited “directory information”). The report recommends that the Census Bureau develop address-update and contact protocols with GQ facilities by type and convene GQ stakeholders in a discussion of access to data and electronic provisions of data for GQ residents.

### **Measurement of Race and Ethnicity**

The quality of the race and ethnicity data collected in the 2020 Census is of central concern; these data are among the most important a U.S. census collects, and the census historically has not counted all racial and ethnic groups equally well. The measurement of race and ethnicity in the 2020 Census was complicated by several factors, including increased capacity for write-in responses and the Census Bureau’s revisions to its coding procedures for reported race and Hispanic origin values. The report concludes that the 2020 Census depicted a more diverse population than did the 2010 Census, but that the 2020 Census had poorer data on race and ethnicity compared with the 2010 Census in terms of coverage error, rates of missing and imputed responses, and the noise infused by new confidentiality protection techniques in 2020.

The report recommends that the Census Bureau conduct research on how changes in format and processing affected the distributions of race and ethnicity. Looking forward to 2030, the report urges the U.S. Office of Management and Budget to finalize its guidance on race and Hispanic origin data collection as soon as practicable — ideally permitting a combined, check-more-than-one race/ethnicity question and allowing the Census Bureau time to test the new question ahead of the 2030 Census. In addition, the report recommends that the Census Bureau consult the redistricting community to determine the optimum set of tabulations to include on the 2030 Redistricting File.

## **IMPACT OF DISCLOSURE AVOIDANCE CHANGES**

### **ON 2020 CENSUS DATA PRODUCTS**

Among the unique challenges confronted by the 2020 Census, the one that is arguably the most consequential and damaging was the result of the Census Bureau's own decisions. The Census Bureau decided in the very late stages of 2020 Census planning to completely replace its Disclosure Avoidance System (DAS) for protecting confidentiality in census data products with an entirely new approach—one that had not been tested, prototyped, or deployed in the population census context. The result has been extensive delays in the production of 2020 Census data relative to previous decades and commensurate weakening of the relevance and accuracy of those data products, to the detriment of the Census Bureau and its user constituency. The stoppage in census field work due to the COVID-19 pandemic understandably explains the delay in producing state-level apportionment totals and the Redistricting Data File compared to their statutory deadlines, but the subsequent delay of many months in release of other data products stems from the lack of disclosure avoidance machinery to produce credible results.

The report firmly states that confidentiality protection is an important principle for the census and for statistical agencies in general and concludes that it is difficult for the census to balance the need for accurate data for small areas and small population groups with an adequate level of protection. However, the report concludes that the Census Bureau's implementation of the new, not-well-developed protection techniques went counter to long-standing principles of decennial census planning. In addition, it is not clear that the chosen privacy budgets for the 2020 Census data products provide much actual protection.

The report recommends that the Census Bureau revisit the risk-utility framework for the 2030 Census to achieve better balance and argues that it should provide as-collected (unperturbed) total population counts for all governmental unit and quasi-governmental units, of all sizes, in the 2030 Census.

For the 2030 Census data product plan, the Census Bureau should begin immediately on a research program with the data user and stakeholder community with a goal of producing a suite of 2030 data products that serve user needs. The Census Bureau should welcome initiatives to add language to appropriate legislative vehicles that prescribes responsibilities and penalties for data users in addition to agency staff for willful, harmful disclosure of confidential information.

### **LEARNING FROM 2020 AND PREPARING FOR 2030**

The report cautions against unduly casting the 2020 Census as the result of a one-off set of conditions; there is still much to be learned from the 2020 experience for shaping the 2030 Census and its successors. The report encourages documenting and improving upon the contingency procedures that had to be developed and invoked extemporaneously in 2020, some of which of which may hold promise for 2030 and beyond.

Looking ahead to the 2030 Census and beyond, the report recommends that the Census Bureau focus its primary attention on a small and manageable number of major innovation areas and pursue a rigorous program of testing and systems development to address those areas.

It is vitally important that the goals and designs of the 2030 Census be developed in true partnership with census data users; community stakeholders; and state, local, tribal, and government partners that make the decennial census the essential, grand civic ceremony upon which the nation relies.

## PANEL TO EVALUATE THE QUALITY OF THE 2020 CENSUS

### TERESA A. SULLIVAN (Chair)

Department of Sociology,  
University of Virginia

### MARGO ANDERSON

Department of History,  
University of Wisconsin–Milwaukee  
(emerita)

### ROBERT M. BELL

Google and AT&T Labs (retired)

### KATHRYN EDIN

Department of Sociology,  
Princeton University

### MARC HAMEL

Statistics Canada (retired)

### GEORGE T. LIGLER

Department of Multidisciplinary  
Engineering, Texas A&M University

### THOMAS A. LOUIS

Department of Biostatistics,  
Johns Hopkins University (emeritus)

### LLOYD B. POTTER

Texas Demographic Center,  
University of Texas at San Antonio

### JOSEPH J. SALVO

University of Virginia  
Biocomplexity Institute

### REGINA SHIH

Department of Epidemiology,  
Rollins School of Public Health,  
Emory University

### C. MATTHEW SNIPP

School of the Humanities and Sciences,  
Stanford University

### EDWARD TELLES

Department of Sociology,  
University of California, Irvine

### WENDY UNDERHILL

National Conference of State  
Legislatures, Denver

### DAVID VAN RIPER

Minnesota Population Center,  
University of Minnesota

## STUDY STAFF

### DANIEL CORK

Study Director

### CONSTANCE F. CITRO

Senior Scholar

### MICHAEL L. COHEN

Senior Program Officer

### ANTHONY MANN

Senior Program Associate

### KATRINA BAUM STONE

Senior Program Officer

## FOR MORE INFORMATION

This Consensus Study Report Highlights is prepared by the Committee on National Statistics based on the Consensus Study Report *Assessing the 2020 Census: Final Report (2023)*.

The study was sponsored by the United States Census Bureau. Any opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of any organization or agency that provided support for the project.

This Consensus Study Report is available from the National Academies Press |  
(800) 624-6242 | <http://www.nap.edu> |  
<http://www.nationalacademies.org>

To read the full report, please visit

<http://www.nationalacademies.org/27150/>

## Division of Behavioral and Social Sciences and Education

**NATIONAL  
ACADEMIES** Sciences  
Engineering  
Medicine

The National Academies provide independent, trustworthy advice that advances solutions to society's most complex challenges.

[www.nationalacademies.org](http://www.nationalacademies.org)