AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics®

FINAL REPORT

SEPTEMBER 2021

2020 Census State Population Totals:

A Report from the American Statistical Association Task Force on 2020 Census Quality Indicators

Table of Contents

Task Force Members*

Nancy A. Potok – Co-Chair
Independent Consultant
Former Chief Statistician of the United States
Former Deputy Director, US Census Bureau

Constance F. Citro
Senior Scholar, National Research Council, Committee on National Statistics
Former Staff Director, National Research Council, Committee on National Statistics

Robert E. Fay
Senior Scientist, Westat
Former Senior Mathematical Statistician, US Census Bureau

Robert Groves
Provost, Georgetown University
Chair, National Research Council, Committee on National Statistics
Former Director, US Census Bureau

Howard Hogan
Adjunct Professor, Department of Statistics, The George Washington University
Former Chief Demographer and Chief of the Decennial Statistical Studies Division, US Census Bureau

D. Sunshine Hillygus
Professor, Duke University, Department of Political Science
Director, Duke Initiative on Survey Methodology
Former Member, Census Scientific Advisory Committee

Thomas A. Louis
Professor Emeritus, Biostatistics, Johns Hopkins Bloomberg School of Public Health
Former Chief Scientist, US Census Bureau

Denice Ross
Fellow, Beeck Center for Social Impact + Innovation, Georgetown University

Matthew Snipp
Burnett C. and Mildred Finley Wohlford Professor of Humanities and Sciences,
Department of Sociology, Stanford University
Director, Institute for Research in the Social Sciences

John H. Thompson
Distinguished Institute Fellow, Biocomplexity Institute and Initiative, University of Virginia
Former Director, US Census Bureau

Katherine K. Wallman
Former Chief Statistician of the United States
1992 President, American Statistical Association

*Robert Santos and Kenneth Prewitt resigned from the task force in April 2021, Santos upon becoming President Biden's nominee to be director of the US Census Bureau and Prewitt upon being named a senior adviser in the office of the director of the US Census Bureau. Some task force members have spoken publicly, testified before Congress, contributed to litigation, written for news outlets, and taken related steps regarding the challenges facing the 2020 Census.

Executive Summary

In October 2020, the American Statistical Association 2020 Census Quality Indicators task force (task force) issued a report on the need to measure the quality of the 2020 Census. Because of the possibility apportionment data would be released on December 31, 2020, or shortly thereafter, the task force promptly brought in a team of outside experts and, with the cooperation of the Census Bureau (the bureau), set out to look at the quality of the 2020 state population totals the bureau ultimately released April 26, 2021.

Below are the conclusions and recommendations of the task force after examining 10 process statistics provided by the outside experts pertaining to the collection and post-data collection operations of the 2020 Census.

Conclusions

1. Indicators released to date by the bureau do not permit a thorough assessment of the 2020 Census data quality. For example, the percentage of completed enumerations in a state is not sufficient to draw conclusions about the quality of the count. Although the task force proposed a variety of possible indicators about census operations that could help evaluate the quality and accuracy of the data more quickly and thoroughly than in previous censuses, only a limited set of indicators has been publicly released thus far, and the task force only had access to state-level indicators composed of process statistics; that is, statistics about the conduct of various 2020 Census operations.
2. Despite concerns that census numbers could be jeopardized by political interference, the task force found no evidence of anything other than an independent and professional enumeration process by the Census Bureau. The bureau appropriately delayed release of data products to ensure careful review and processing of the data according to bureau quality standards.
3. Across the limited set of state-level process statistics evaluated by the task force, it found no major anomalies that would indicate census numbers are not fit for use for purposes of apportionment.
4. The task force’s ability to more thoroughly evaluate the quality, accuracy, and coverage of the 2020 Census has been hampered by limits on available information and research. Many of the proposed indicators could not be calculated because household characteristics data were not yet available. For several of the process statistics calculated, the task force has not found research that provides a clear enough understanding of their implications for erroneous or incomplete enumerations.

5. The set of process statistics evaluated by the task force is relevant for evaluating the quality of census numbers for apportionment, but not redistricting or distribution of federal funds. These and other assessments require scrutiny at more detailed levels of geography and subgroups of the population.

Recommendations

1. The Census Bureau is to be commended for entering into work with the National Academy of Sciences Committee on National Statistics (CNSTAT) and other experts to conduct a more thorough assessment of 2020 Census data quality. This evaluation should expand the set of process statistics the task force's outside experts accessed and examine patterns at more detailed levels of geography (e.g., census tracts) and for population subgroups. The evaluation should scrutinize the increase in missing household characteristics, new procedures for counting overseas population, late-breaking changes in methods for using administrative records to enumerate nonresponding households, increased uses of imputation including for group quarters, and prompt determination of any increase in undercount of Black and Hispanic populations and children relative to 2010 based on demographic analysis.
2. Planning for the 2030 Census should incorporate explicit attention to evaluating and reporting on data quality. In particular, the Census Bureau needs to invest in the research necessary to enable the use of process statistics for evaluating data quality before apportionment counts are released for future censuses. The historic approach to census quality assessments (e.g., coverage assessments based on the post-enumeration survey) happens long after the release of census data products, but this is no longer sufficient since process statistics and other data can be available much sooner.

The 2020 Census was conducted during a "perfect storm" of a pandemic, weather disasters, and attempts at political interference—it was a census unlike any other in American history. In this context, for the state-level data released on April 26, 2021, the task force focused on 10 measures of census operations that contribute to a successful enumeration. The task force was granted access, via its small team of sworn-in experts, to the bureau's files used to produce state population totals, and they began receiving tabulations of candidate process indicators serially beginning in the spring of 2021. The task force experts team has not received or examined any non-public tabulations lower than the state level nor tabulations with any characteristics. The task force examined 10 state-level statistics on census processes that could indicate potential problems with total population counts. This effort was always to be a stopgap until more complete analyses could be done by others over a longer time with more data. Because the bureau has entered into an

agreement with CNSTAT to continue the work of assessing the quality of the 2020 Census, this will be the last report of the task force.

Acknowledgements

The task force thanks the Census Bureau; task force members Robert Fay and Howard Hogan; and the team of Paul Biemer, Joseph Salvo, and Jonathan Auerbach for their extensive contributions that provided invaluable input for this report.

The task force concludes that the 10 process statistics it examined at the state level provide insufficient information for a summary judgment about the quality of the 2020 Census and recommends two high priority lines of research on data released later than the state-level counts: one that examines these process statistics at more detailed levels of geography and another that examines the Census Bureau's traditional methods for assessing under- and over-counts.

Introduction

In October 2020, the American Statistical Association 2020 Census Quality Indicators task force (task force) issued a report with a series of recommendations regarding the need to measure the quality of the 2020 Census. The task force described indicators in a general fashion with several examples to illustrate their intended use, while stating that a more detailed set of quality indicators should be identified. The indicators included self-response data to identify any lack of uniformity across population groups and geographic entities such as states, cities, counties, and rural areas, which would indicate inequity of the counts across the US and field data collection and post-collection data processing quality measures that can be calculated for each of the field and processing operations. In the October 2020 report, the task force also recommended the following:

1. The indicators should be readily available and used expeditiously to assess the quality of the 2020 Census.
2. Qualified external researchers should be granted access to the data to help conduct the analyses.
3. Additional assessments should be conducted when more data become available.
4. Early planning for the 2030 Census should build on the lessons of 2020, be conducted in public, and include extensive stakeholder input.

5. The Census Bureau's authorizing statute, Title 13 US Code, should be updated. Title 13 needs to align better with the more recent Information Quality Act and OMB guidelines, and it should be determined if further amendments to protect the integrity and independence of the Census Bureau and decennial census are warranted.

This report addresses recommendations 1, 2, and 3 as they relate to state population totals used for apportionment. The task force asked a team of researchers to work with the bureau to produce materials relevant to the goals laid out by the task force. The result was a series of process statistics from the experts that the task force has examined and is making available in our report. In addition, two task force members conducted work to expand on relevant areas considered important to future research: administrative records and demographic benchmarks. The work of all the researchers is available at the sites referenced in this document.

The obstacles to the 2020 Census caused concern among stakeholders that the accuracy of the 2020 Census enumeration could be impaired even with the best efforts of bureau staff. The task force commends the bureau for conducting 2020 Census operations in a professional manner and taking the necessary time to review the data and implement corrective actions as necessary. Despite concerns that the apportionment counts would be jeopardized by political interference, the final census numbers were not politicized.

The 10 process statistics in this report span census operations, including Master Address File (MAF) development, self-response, nonresponse follow-up (NRFU), data compilation and processing, and group quarters enumeration. Examination of these statistics suggests potential lines of research, and they may prove useful to knowledgeable data users in particular states. However, they are difficult to interpret on their own: Some relate to processes that cannot be compared with 2010 (e.g., self-responses without a MAF ID), while others show little variation across states (e.g., duplicate responses). For some statistics, prior research finds a high propensity for erroneous and incomplete enumerations (e.g., NRFU interviews from a landlord or other proxy), but the extent of erroneous enumerations or omitted people is unknown for others (e.g., administrative records).

For the above reasons, we do not draw conclusions about the quality of the 2020 Census population totals for states from these statistics. We further recommend against combining them into a single process indicator by state—this cannot be supported by science as we understand it.

Background

The task force report was based on three sources: process statistics for key 2020 Census operations; a review of the research supporting the use of administrative records in the

2020 Census; and an examination of demographic benchmarks, contrasting the 2020 Census with previous censuses.

Process Statistics

Biemer et al.[1] focused on specific 2020 Census data collection operations briefly described as follows:

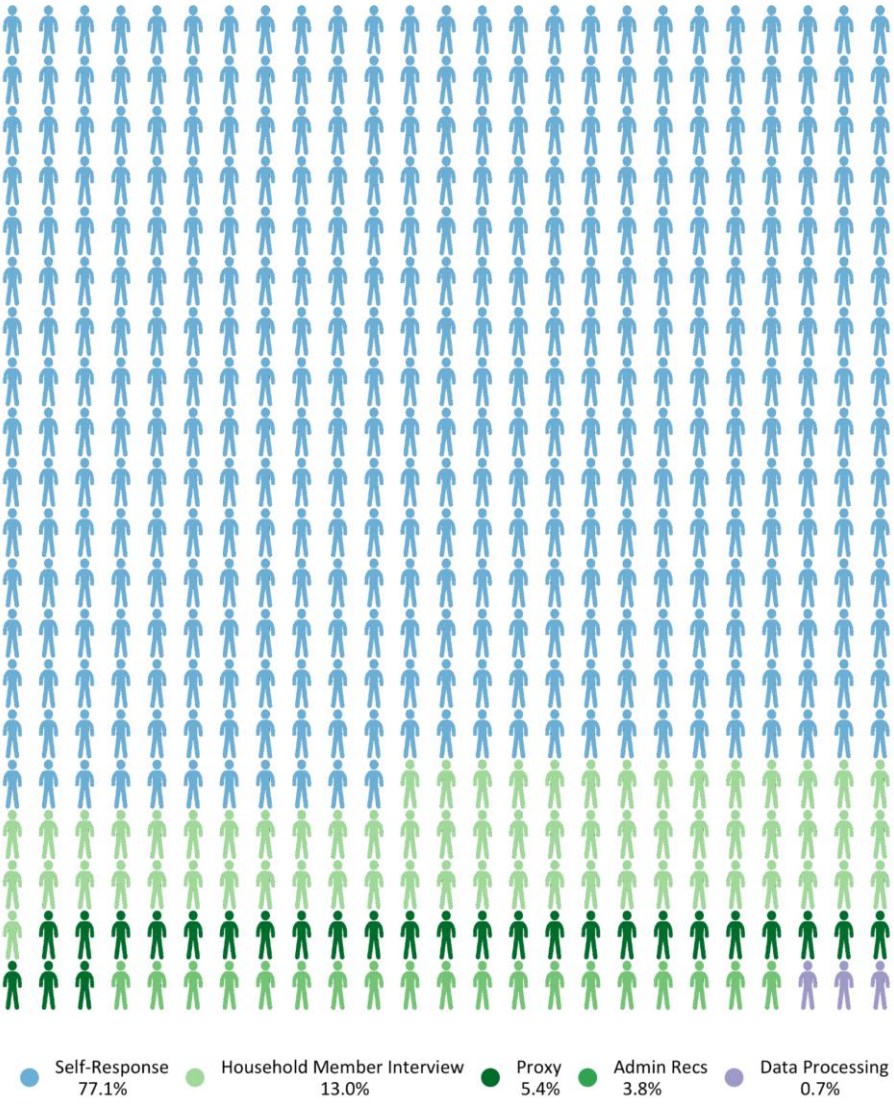
1. Master Address File (MAF) Development – Update the initial list of residential addresses used by the Census Bureau to contact households and elicit their response to the census.
2. Self-Response – Encourage a household resident to complete the questionnaire by mail, telephone, or internet.
3. Nonresponse Follow-Up (NRFU) – Each address from which a self-response was not received is visited at least once by a census enumerator. The outcomes of the NRFU enumeration fall into the following six broad categories:
    a. The address is determined to be for a vacant or nonexistent housing unit.
    b. The address is determined to be occupied and enumeration is obtained through self-response with a household member.
    c. The address is determined to be an occupied housing unit and enumeration is obtained by administrative records after at least one visit by an enumerator.
    d. The housing unit is determined to be occupied and enumeration is obtained through a proxy respondent.
    e. The housing unit is determined to be occupied and only minimal information, such as only a population count, is obtained.
    f. The status of the address cannot be determined, resulting in a status of unresolved.
4. Data Compilation and Processing – The results of the data collection operations are processed through several steps that produce the Census Unedited File (CUF), which is used to provide the apportionment counts. These processes include the use of statistical methods to fill in missing information and resolve multiple responses from the same address. For a subset of the NRFU unresolved addresses, the statistical processes will result in a status of occupied and an assignment of population count.
5. Group Quarters – Enumerate persons in group facilities such as student housing, nursing homes, and military barracks.

---

[1] Paul Biemer, Joseph Salvo, and Jonathan Auerbach, *The Quality of the 2020 Census: An Independent Assessment of Census Bureau Activities Critical to Data Quality*, https://www.amstat.org/asa/files/pdfs/POL-BiemerSalvoAuerbachReport.pdf.

Figure 1 below describes how the occupied housing units were enumerated in the 2020 Census. Each person represents approximately 250,000 occupied housing units color coded by how they were enumerated (Biemer et al.).

Figure 1. How Occupied Households Were Enumerated in the 2020 Census



Biemer et al. produced 10 process statistics based on information provided by the Census Bureau for components of the 2020 Census operations described above. Table 1 summarizes these process statistics.

Table 1. Summary of Process Statistics

Process Statistics*		Description
MAF Development (MAF)		
1	MAF Revisions	Percent of all addresses that were either deleted or added during the 2020 census data collection period
Self-Response (SR)		
2	Questionnaires without identification (ID) not on MAF (Non-Matching No IDs)	Percent of housing units (HUs) submitting questionnaires without census IDs and no matching address was found on the MAF for 2020
3	Multiple Responses	Percent of occupied HUs that submitted two or more questionnaires for 2020 minus the corresponding percentage for 2010
4	Usual Residence at College (URC)	Percent of occupied HUs with two or more people where one or more occupant indicated their usual residence was at college for 2020 minus the corresponding percentage for 2010
Nonresponse Follow-up (NRFU)		
5	Responses Obtained by Proxy (Proxy)	Percent of persons in occupied HUs whose count was obtained by proxy interview for 2020 minus the corresponding percentage for 2010
6	Enumerations with Only a Population Count (Count Only)	Percent of occupied HUs where only a population count was obtained for 2020 minus the corresponding percentage for 2010
7	Enumerations via Administrative Records (Admin Recs)	Percent of occupied HUs enumerated by administrative records for 2020
Data Processing		
8	MAF Addresses Having Imputed Status (Status Imputation)	Percent of MAF units whose status was imputed for 2020 minus the corresponding percentage for 2010
9	Occupied Housing Units with Imputed Population Counts (Count Imputation)	Percent of occupied HUs with known status but whose population count was imputed for 2020 minus the corresponding percentage for 2010
Group Quarters (GQ)		
10	Group Quarters with Imputed Count (GQ Imputation)	Percent of the GQ population that was imputed in 2020

*All data were approved for public dissemination by the Census Bureau Disclosure Review Board. Clearance numbers: CBDRB-FY21-DSSD007-0021, CBDRB-FY21-DSSD007-0024, CBDRB-FY21-DSSD007-0026.

Review of Research Supporting the Use of Administrative Record Enumeration in the 2020 Census

There is an extensive body of research in the public domain that documents the development of the administrative records (AR) program employed in the 2020 Census. Robert Fay prepared a summary of this research[2] with a focus on two questions: (1) To the extent AR enumerations were incorporated into the 2020 Census, did they maintain the accuracy that would have been obtained by using another approach? (2) Is there currently enough information to answer the first question? Dr. Fay's report is discussed in the Findings section.

Demographic Benchmarks

The Census Bureau produces two demographic benchmarks to which the Census can be compared. These are known as the demographic analysis (DA) estimates and the population estimates. Both sets rely on measures of births, deaths, and migration to construct alternative population estimates. They differ principally in their baseline starting points. The population estimates start with the previous census and project that forward 10 years to the time of the current census. The demographic analysis estimates use as their base the past 65 years of birth records plus information from administrative records to construct a set of estimates of the population at the time of the current census. Because of this long data series, demographic analysis only produces estimates at the national level. The population estimates track internal migration over the 10 years and thus can produce sub-national estimates. Howard Hogan prepared a summary of these demographic benchmarks[3] contrasting the results from the 2020 and previous censuses. Dr. Hogan's report is discussed in the Findings section below.

Results

Biemer et al. worked closely with Census Bureau staff to produce the process statistics described above. This was a lengthy, iterative process, and the task force commends the exceptional effort required to produce these data. The process statistics Biemer et al. developed are shown below.

---

[2] Robert E. Fay, *Review of Research Supporting the Use of Administrative Record Enumeration in the 2020 Census*, https://www.amstat.org/asa/files/pdfs/POL-Fay-CQI-Report.pdf.

[3] Howard Hogan, *A Note on Demographic Benchmarks,* https://www.amstat.org/asa/files/pdfs/POL-Hogan-CQI-Report.pdf.

Table 2 presents the process statistics calculated for each state. In some cases (e.g., responses obtained by proxy), the process statistics represent a comparison between the 2020 and 2010 censuses. In other cases (e.g., MAF revisions), the process statistics present results solely for the 2020 Census.

Table 2. Process Statistics by State

State	Relative Difference between the Census and the Population Estimates	1. MAF Revisions (2020 Only)	2. Questionnaires without ID not on MAF (2020 Only)	3. Multiple Responses (2020 minus 2010)	4. Usual Residence at College (2020 minus 2010)	5. Responses Obtained by Proxy (2020 minus 2010)	6. Enumerations with Only a Population Count (2020 minus 2010)	7. Enumerations via Administrative Records (2020 Only)	8. MAF Addresses Having Imputed Status (2020 minus 2010)	9. Occupied Housing Units with Imputed Population Counts (2020 minus 2010)	10. Group Quarters in TEA 1 & 6 with Imputed Count (2020 Only)
US	0.62	9.22	8.79	17.53	0.53	-0.35	0.53	3.84	0.74	-0.22	2.01
AL	2.11	14.38	7.21	15.18	0.6	-1.72	-0.26	3.51	0.9	-0.44	4.98
AK	0.18	26.55	21.55	20.14	0.18	-0.62	1.3	2.11	0.81	-0.25	4.76
AZ	-3.28	10.65	11.25	18.13	0.35	-0.79	-1.02	3.41	0.48	-0.33	3.19
AR	-0.6	16.01	7.52	16.69	0.2	-1.37	0.54	3.36	0.71	-0.17	3.57
CA	0.3	8.82	6.47	18.82	0.78	-0.11	0.79	3.75	0.72	-0.07	2
CO	-0.42	9.95	8.16	18.09	0.54	-1.11	-0.84	3.55	0.66	-0.44	1.1
CT	1.25	7.56	11.58	17.41	0.66	0.21	1.06	4.27	0.84	0.01	0.5
DE	0.51	6.24	5.38	16.68	0.62	-1.78	-0.08	4.87	0.25	0.02	11.26
DC	-3.18	11.51	7.16	16.52	-0.11	-0.98	1.26	4.98	0.74	-0.67	0.26
FL	-0.69	7.91	7.9	16.97	0.42	-0.28	0.15	4.43	0.64	-0.38	3.53
GA	0.15	12.18	5.88	17.88	0.68	-0.87	-0.61	4.32	0.71	-0.6	1.99
HI	3.17	14.28	14.09	15.66	0.67	-0.01	-1.25	1.65	1	-0.14	0.63
ID	1.15	11.17	8.53	15.95	0.13	-0.64	-0.38	2.51	0.53	-0.27	0.5
IL	1.56	8.71	12.03	17.04	0.64	0.09	1.1	3.64	0.73	-0.15	1.04
IN	0.49	7.25	5.51	15.43	0.37	-0.21	0.91	3.62	0.6	-0.01	0.64
IA	0.83	7.27	5.14	17.21	0.76	0.29	1.09	4.23	0.54	-0.01	2.25
KS	0.78	7.98	4.95	16.04	0.59	0.69	1.2	3.05	0.66	0.01	0.57
KY	0.62	11.91	7.93	15.45	0.23	-1.74	0.13	3.18	0.69	-0.08	2.17
LA	0.15	12.77	6.45	17.2	0.28	-1.9	-0.19	5.12	1.36	-0.49	4.34
ME	0.94	11.9	7.79	14.98	0.32	0.03	0.85	4.12	0.8	-0.05	5.16
MD	1.94	5.87	4.28	17.78	0.55	-0.31	0.79	4.04	0.55	-0.18	1.23
MA	1.91	10.18	10.6	17.46	0.45	0	1.18	4.19	1.06	0.05	0.58
MI	1.01	7.89	5.5	17.39	0.76	-0.87	0.51	3.36	0.69	-0.12	4.81
MN	0.9	6.48	6.08	18.34	0.79	-0.03	0.95	2.85	0.6	-0.11	0.32
MS	-0.34	14.23	7.25	15.25	0.2	-1.89	0.87	3.89	0.84	-0.47	6.81
MO	0.05	13.19	6.46	16.24	0.64	-0.91	0.62	3.61	0.75	-0.12	3.95
MT	0.54	19.3	12.3	19.88	0.37	-0.42	-0.6	2.73	0.56	-0.1	2.01
NE	1.25	7.79	5.85	15.73	0.52	-0.07	0.68	3.47	0.62	-0.04	0.41
NV	-0.76	7.79	6.63	18.97	0.25	-1.73	-1.22	4.39	0.55	-0.41	1.72
NH	0.88	8.79	7.45	18.87	0.55	-0.35	0.88	4.46	0.69	0	0.13
NJ	4.48	8.33	12.71	18.52	0.58	0.14	1.37	4.09	0.78	-0.28	0.68
NM	0.54	19.68	16.42	20.15	0.39	-1.7	-0.62	2.68	0.71	-0.71	3.15
NY	4.23	13.32	15.76	19.61	0.21	-0.29	1.59	4.37	1.19	-0.23	0.79
NC	-1.35	10.55	5.29	17.35	0.58	-1.17	-0.11	4.52	0.67	-0.43	3.07
ND	1.81	14.98	11.02	16.76	0.39	0.28	1.15	2.9	0.77	-0.07	0.32
OH	0.85	7.19	6.1	16.97	0.56	-0.07	0.7	4.07	0.65	-0.06	1.14
OK	-0.46	15	9.71	16.91	0.23	-0.42	0.63	2.85	0.65	-0.05	0.53
OR	0	10.2	7.53	17.23	0.44	-0.11	0.8	2.97	0.63	-0.24	1.64
PA	1.63	8.83	8.76	15.9	0.62	0.34	0.93	4.04	0.73	-0.08	1.6
RI	3.72	7.83	15.49	17.1	0.17	1.74	0.8	5.09	1.03	0.01	0.21
SC	-1.68	12.56	4.9	15.38	0.25	-0.28	0.3	4.26	0.69	-0.33	3.49
SD	-0.56	12.89	19.2	15.94	0.51	-0.37	0.29	3.1	0.63	-0.34	1
TN	0.51	10.29	5.15	15.7	0.43	-1.07	0.35	3.94	0.67	-0.32	3.62
TX	-0.48	11.25	6.97	19.04	0.59	0.07	0.06	3.97	0.7	-0.39	1.9
UT	0.99	11.01	13.59	18.32	0.17	0.77	1.13	2.54	0.8	-0.03	0.55
VT	3.09	13.63	9.24	16.67	0.4	0.69	1.41	3.09	0.85	0.05	0.7
VA	0.51	6.75	4.8	17.65	0.89	-0.49	0.13	3.64	0.6	-0.13	0.71
WA	0.35	7.8	8.37	16.6	0.47	-0.38	0.88	3.25	0.81	-0.2	2.72
WV	0.31	25.58	15.07	16.52	0.13	-2.11	0.62	2.74	0.75	0.02	0.53
WI	1.03	7.07	6.81	16.31	0.77	0.21	1.01	3.17	0.67	-0.09	1.23
WY	-0.89	15.36	10.68	16.55	0.31	0.24	0.04	3.38	0.67	-0.31	1.92

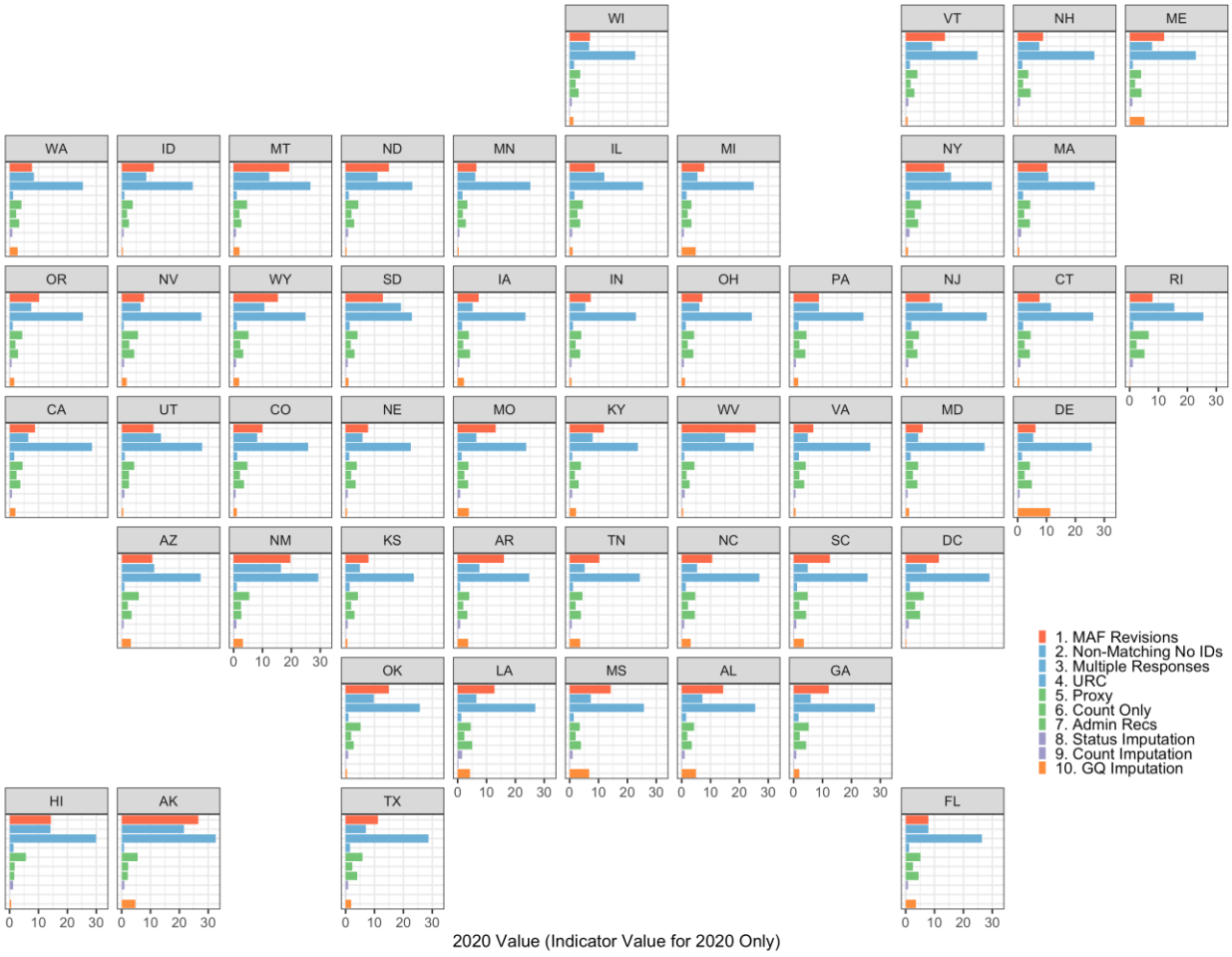
Table 3 presents the 2020 Census components of the process statistics that compared the 2020 and 2010 censuses. For example, the response by proxy statistic subtracted the percentage of total population obtained from proxy responses in 2010 from the corresponding percentage of proxy responses in 2020. The data in Table 3 show the percentage of persons whose response was obtained by proxy in 2020.

Table 3. 2020 Components for 2020 - 2010 Process Statistics

State	3. Percent of Occupied HUs Having Multiple Responses	4. Percent of Occupied HUs with Usual Residence at College	5. Percent of Persons whose Responses were Obtained by Proxy	6. Percent of Occupied HUs with Only a Population Count	8. Percent of MAF Addresses with Imputed Status	9. Percent of Occupied HUs with Imputed Population Counts
US	26.47	1.45	4.65	2.28	0.88	0.06
AL	25.44	1.51	4.34	2.01	1.09	0.1
AK	32.55	0.84	5.52	2.29	1.01	0.05
AZ	27.39	0.95	5.87	2.1	0.68	0.03
AR	24.81	0.9	3.99	1.89	0.91	0.05
CA	28.56	1.62	4.42	2.44	0.81	0.05
CO	25.83	1.24	4.71	2.18	0.77	0.06
CT	26.15	1.87	4.48	2.35	0.98	0.05
DE	25.65	1.54	4.29	2.38	0.64	0.08
DC	29.06	1.4	6.34	3.31	1.01	0.1
FL	26.39	1.19	5.17	2.46	0.81	0.07
GA	28.15	1.76	5.28	2.11	0.9	0.12
HI	29.96	1.36	5.66	1.69	1.16	0.04
ID	24.6	0.98	3.79	1.89	0.71	0.06
IL	25.5	1.71	4.56	2.69	0.85	0.05
IN	22.92	1.17	4.04	2.04	0.7	0.04
IA	23.43	1.53	3.95	1.87	0.61	0.05
KS	23.56	1.34	4.3	1.85	0.74	0.06
KY	23.58	0.93	3.94	1.87	0.9	0.08
LA	26.84	1.26	4.56	2.27	1.53	0.17
ME	22.87	1.15	4.01	1.85	0.93	0.04
MD	27.31	1.74	4.3	2.44	0.71	0.07
MA	26.71	1.81	4.4	2.27	1.17	0.11
MI	24.98	1.68	3.43	2.08	0.81	0.05
MN	25.17	1.67	3.4	1.83	0.68	0.03
MS	25.7	1.44	3.46	2.09	1.04	0.12
MO	23.76	1.38	3.74	2.26	0.86	0.07
MT	26.54	1.06	4.62	2.03	0.82	0.08
NE	22.49	1.32	3.89	1.95	0.71	0.02
NV	27.57	0.74	5.69	2.65	0.78	0.09
NH	26.56	1.64	3.7	1.91	0.8	0.05
NJ	28.18	1.93	4.59	2.6	0.92	0.06
NM	29.31	1.05	5.47	2.51	0.99	0.06
NY	29.86	1.46	5.44	3.17	1.39	0.09
NC	26.85	1.46	4.71	2.19	0.89	0.09
ND	23.13	1.18	4.4	2.03	0.89	0.03
OH	24.32	1.4	4.33	2.08	0.73	0.05
OK	25.66	1	5.23	1.87	0.82	0.03
OR	25.33	1.08	4.36	2	0.74	0.06
PA	24.18	1.73	4.44	2.03	0.86	0.07
RI	25.57	1.29	6.65	2.46	1.15	0.1
SC	25.59	1.2	4.79	1.94	0.91	0.08
SD	22.92	1.37	4.17	1.77	0.79	0.03
TN	24.26	1.21	4.44	2.12	0.8	0.06
TX	28.64	1.46	5.84	2.34	0.85	0.05
UT	27.9	1.12	4.35	2.52	0.9	0.05
VT	24.87	1.41	4.08	1.79	0.99	0.06
VA	26.58	1.82	4.17	2.01	0.72	0.04
WA	25.41	1.15	4.04	2.23	0.93	0.04
WV	24.96	0.95	4.44	1.69	0.99	0.07
WI	22.72	1.51	3.64	2.04	0.78	0.05
WY	24.83	1.09	5.16	2.25	0.89	0.02

Figure 2 presents a graphical display by state of the Table 3 results combined with the process statistics only available for the 2020 Census.

Figure 2. State Results for Process Statistics



Four of the process statistics were only available for the 2020 Census, and six compare the 2020 and 2010 censuses. As indicated in tables 2 and 3, the four process statistics available only for the 2020 Census are (1) MAF revisions, (2) questionnaires without IDs, not on the MAF, (3) enumeration with administrative records, and (4) group quarters in type of enumeration areas (TEAs) 1 and 6 with imputed counts.

The four process statistics available only for the 2020 Census are interesting, but the task force did not find any research to assess to what degree they represent potential for error. These four process statistics represent operations intended to maintain or improve the quality of the 2020 Census. It could well be states with high scores for one or more of these process statistics have higher quality data and less risk for error than states with low scores.

The six process statistics that compare the 2020 and 2010 censuses represent situations in which errors—both undercounts and overcounts—may occur. However, a number of important questions, many of which Biemer et al. identified, must be answered before these process statistics can be associated with any specific level of error risk. For example, the 2020 Census processes differed dramatically from those used in the 2010 Census. Therefore, it is critical that differences be further examined to determine the extent to which they simply reflect improvements to census-taking methodology.

Biemer et al. presented a detailed description of the process statistics to the task force. They also conducted a preliminary analysis of these data. Biemer et al. stated the process statistics should not be interpreted as error rates and concluded that, given the data at hand, their review did not find conclusive evidence that state-level counts used for apportionment purposes are of lower quality in 2020 than in 2010. Nor is there evidence that the apportionment count for any given state is in error. However, the absence of evidence revealed by the process statistics is also insufficient to state conclusively that there were no errors in the apportionment counts. Biemer et al. recommended important future research that should be conducted to better understand each of the process statistics.

Biemer et al. presented one approach to how the process statistics might be used to assess the 2020 Census. This approach determined each of the process statistics represented an error risk (not actual error, just the potential for error being introduced into the population counts). Using that approach, Biemer et al. ranked the states based on each of the process statistics, implying that some states had a higher risk of error than others. In addition, Biemer et al. combined the 10 process statistics into a summary process statistic to represent the overall risk of error for each state. Biemer et al. then ranked the states based on the summary process statistic.

The task force conducted an extensive review of the report prepared by Biemer et al. and found it to be an important first step in understanding the quality of the 2020 Census. The task force agrees there is no conclusive evidence that state-level counts used for apportionment are of lower quality in 2020 relative to 2010, and there is no evidence the apportionment count in any given state is in error. The task force also agrees that the development of the process statistics is an important first step in understanding the results of the 2020 Census and further research is required.

However, the task force finds it is premature to use the individual process statistics as direct measures of error risk or combine them into a summary process statistic. In addition, it is also premature to use the individual process statistics as a basis for ranking states. At this point, there remains much to be learned about how the specific process statistics relate to census errors; that is, whether they indicate a greater or lesser risk of net error.

Finally, the task force notes the development of a summary or composite statistic that can be used to understand the overall quality of the state apportionment counts is an important concept for future research. However, the task force concludes there is insufficient information at this time to support analysis based on the current summary process statistic proposed by Biemer et al. To properly interpret any summary of individual statistics, it is first necessary to understand how the statistics interact. For example, do some process statistics overlap with others so the effects they measure are double counted or cancel out in a composite summary? Additionally, the concerns described above regarding the individual process statistics are heightened when they are combined into a composite. Currently, the information to interpret the results of the summary process statistic does not exist. Given this lack of information and insight, the task force strongly cautions against using the individual and summary process statistics to assess the potential for errors in the census until additional research can be conducted with more detailed data at lower levels of geography than state totals.

Review of Research Supporting the Use of Administrative Record Enumeration in the 2020 Census

Robert Fay's report documents a large body of research carried out by the Census Bureau in developing the 2020 Census administrative records program. Fay concludes several considerations are sufficient reason to withhold judgment on whether the administrative record (AR) enumerations in the 2020 Census were of equal quality to the NRFU interviews they replaced. The 2020 Post Enumeration Survey (PES), in a role like the 2010 Census Coverage Measurement (CCM), could clarify the accuracy achieved by AR enumerations, depending on the success of the PES effort. His report includes the following observations and recommendations:

> Some arguments support the quality of AR enumerations as implemented in the 2020 Census as a replacement for NRFU enumeration:
>
> - In Brown, Childs, and O'Hara (2015) the accuracy of AR enumeration approached that of NRFU at the upper end of the predicted agreement rate of AR with the population count.
>
> - Particularly after the 2016 test census, a number of modifications have been introduced to address specific concerns that have been identified and documented. The cumulative effect of these improvements potentially has raised the quality of the AR enumerations implemented in 2020 to their expected NRFU level. Until the 2020 PES becomes available, there was no obvious way to evaluate most of these improvements given time constraints.

- The AR enumerations freed up resources that may have contributed to increasing the completeness of census operations for other NRFU housing units.

A number of concerns can also be mentioned:

- Comparisons of AR enumeration as a substitute for NRFU were not initially favorable.
- Over the span of studies in the decade, the evidence for the quality of AR enumeration relative to NRFU procedures became less quantitative.
- Using 2010 NRFU as training data for the 2020 application may produce suboptimal predictions.
- There were and are opportunities for the Census Bureau to clarify some technical details of the implementation, such as the numerical values of the thresholds.
- Few (perhaps only one) studies examined the impact of AR enumeration on historically undercounted groups.

Demographic Benchmarks

As noted above, Howard Hogan examined demographic analysis (DA) and the population estimates and compared the 2020 benchmarks with previous censuses.

For demographic analysis, Hogan provided the following table comparing net undercount rates for the 2020 and 2010 censuses:

Table 4. 2010 and 2020 Demographic Analysis

	Low	Low Middle	Middle	High Middle	High
2010	1.0	0.4	0.1	-0.4	-1.3
2020		0.2	-0.3	-1.2	

Hogan concluded that if one accepts, as the bureau argues, that the 2010 Low Middle and High Middle are the most comparable to the 2020 estimates, some conclusions are clear. A plausible estimate is that both censuses produced a small overcount. It is equally plausible that the 2010 Census produced a small undercount, while the 2020 Census produced a somewhat larger undercount.

With respect to the Population Estimates, Hogan compared the 2020 Census with the 2010, 2000, and 1990 censuses. Hogan noted the postcensal estimates are compared with the next census at the end of each decade. The difference is termed the Error of Closure (EoC) and has traditionally been used to assess the quality of the postcensal estimates. The difference between the two sets of population measures is comprised of the following:

- Errors in the previous census
- Errors in the current census
- Errors in measuring intercensal change
- Definitional changes between censuses

If there are no errors in the population estimates, the EoC measures the relative coverage between the two censuses. If the two censuses have equal coverage, the EoC measures errors in estimating population change. Obviously, in almost all cases, the EoC results from both changes in census coverage and errors in measuring population change. Table 5 shows National Errors of Closure from 1990-2020

Table 5. National Errors of Closure: 1990 to 2020

Year	EoC
2020	0.62
2010	0.10
2000	2.48
1990	-0.60

Hogan concluded the 2020 Census came in slightly higher than the population estimates, although well within the range observed in recent censuses. The outlier is the 2000 Census, where the population estimates badly underestimated international migration.

Conclusions and Recommendations

Conclusions

1. Indicators released to date by the bureau do not permit a thorough assessment of the 2020 Census data quality. For example, the percentage of completed enumerations in a state is not sufficient to draw conclusions about the quality of the count. Although the task force proposed a variety of possible indicators about census operations that could help evaluate the quality and accuracy of the data more quickly and thoroughly than in previous censuses, only a limited set of indicators has been publicly released thus far, and the task force only had access to state-level indicators composed of process statistics; that is, statistics about the conduct of various 2020 Census operations.
2. Despite concerns that census numbers could be jeopardized by political interference, the task force found no evidence of anything other than an independent and professional enumeration process by the Census Bureau. The bureau appropriately delayed release of data products to ensure careful review and processing of the data according to bureau quality standards.
3. Across the limited set of state-level process statistics evaluated by the task force, it found no major anomalies that would indicate census numbers are not fit for use for purposes of apportionment.
4. The task force's ability to more thoroughly evaluate the quality, accuracy, and coverage of the 2020 Census has been hampered by limits on available information and research. Many of the proposed indicators could not be calculated because household characteristics data were not yet available. For several of the process statistics calculated, the task force has not found research that provides a clear enough understanding of their implications for erroneous or incomplete enumerations.
5. The set of process statistics evaluated by the task force is relevant for evaluating the quality of census numbers for apportionment, but not redistricting or distribution of federal funds. These and other assessments require scrutiny at more detailed levels of geography and subgroups of the population.

Recommendations

1. The Census Bureau is to be commended for entering into work with the National Academy of Sciences Committee on National Statistics (CNSTAT) and other experts to conduct a more thorough assessment of 2020 Census data quality. This evaluation should expand the set of process statistics the task force's outside experts accessed and examine patterns at more detailed levels of geography (e.g., census tracts) and for population subgroups. The evaluation should scrutinize the increase in missing household characteristics, new procedures for counting overseas population, late-breaking changes in methods for using administrative records to enumerate nonresponding households, increased uses of imputation including for group quarters, and prompt determination of any increase in undercount of Black and Hispanic populations, and children relative to 2010 based on demographic analysis.
2. Planning for the 2030 Census should incorporate explicit attention to evaluating and reporting on data quality. In particular the Census Bureau needs to invest in the research necessary to enable the use of process statistics for evaluating data quality before apportionment counts are released for future censuses. The historic approach to census quality assessments (e.g., coverage assessments based on the post-enumeration survey) happens long after the release of census data products, but this is no longer sufficient since process statistics and other data can be available much sooner.