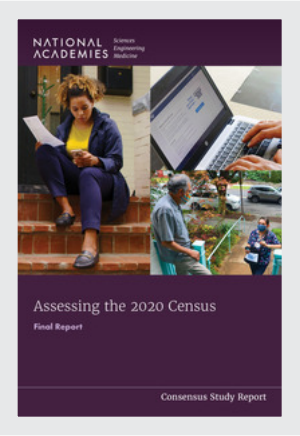
NATIONAL ACADEMIES Sciences Engineering Medicine

NATIONAL ACADEMIES PRESS Washington, DC

This PDF is available at http://nap.nationalacademies.org/27150



Assessing the 2020 Census: Final Report (2023)

BUY THIS BOOK

FIND RELATED TITLES

DETAILS

532 pages | 6 x 9 | PAPERBACK

ISBN 978-0-309-70646-9 | DOI 10.17226/27150

CONTRIBUTORS

Teresa A. Sullivan and Daniel L. Cork, Editors; Panel to Evaluate the Quality of the 2020 Census; Committee on National Statistics; Division of Behavioral and Social Sciences and Education; National Academies of Sciences, Engineering, and Medicine

SUGGESTED CITATION

National Academies of Sciences, Engineering, and Medicine. 2023. *Assessing the 2020 Census: Final Report*. Washington, DC: The National Academies Press. https://doi.org/10.17226/27150.

Visit the National Academies Press at nap.edu and login or register to get:

- Access to free PDF downloads of thousands of publications
- 10% off the price of print publications
- Email or social media notifications of new titles related to your interests
- Special offers and discounts

All downloadable National Academies titles are free to be used for personal and/or non-commercial academic use. Users may also freely post links to our titles on this website; non-commercial academic users are encouraged to link to the version on this website rather than distribute a downloaded PDF to ensure that all users are accessing the latest authoritative version of the work. All other uses require written permission. (Request Permission)

This PDF is protected by copyright and owned by the National Academy of Sciences; unless otherwise indicated, the National Academy of Sciences retains copyright to all materials in this PDF with all rights reserved.

NATIONAL ACADEMIES Sciences Engineering Medicine

NATIONAL ACADEMIES PRESS
Washington, DC

Assessing the 2020 Census

Final Report

Teresa A. Sullivan and Daniel L. Cork,
Editors

Panel to Evaluate the Quality of the 2020 Census

Committee on National Statistics

Division of Behavioral and Social Sciences and Education

Consensus Study Report

Copyright National Academy of Sciences. All rights reserved.

NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

This activity was supported by a contract between the National Academy of Sciences and the U.S. Bureau of the Census (1333LB21D00000003/ 1333LB21F00000121). Support of the work of the Committee on National Statistics is provided by a consortium of federal agencies through a grant from the National Science Foundation (No. 1560294) and several individual contracts. 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.

International Standard Book Number-13: 978-0-309-70646-9
International Standard Book Number-10: 0-309-70646-7
Digital Object Identifier: https://doi.org/10.17226/27150
Library of Congress Control Number: 2023947244

This publication is available from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Copyright 2023 by the National Academy of Sciences. National Academies of Sciences, Engineering, and Medicine and National Academies Press and the graphical logos for each are all trademarks of the National Academy of Sciences. All rights reserved.

Printed in the United States of America.

Suggested citation: National Academies of Sciences, Engineering, and Medicine. 2023. *Assessing the 2020 Census: Final Report.* Washington, DC: The National Academies Press. https://doi.org/10.17226/27150.

Copyright National Academy of Sciences. All rights reserved.

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, nongovernmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Marcia McNutt is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. John L. Anderson is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The National Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at www.nationalacademies.org.

Copyright National Academy of Sciences. All rights reserved.

Consensus Study Reports published by the National Academies of Sciences, Engineering, and Medicine document the evidence-based consensus on the study's statement of task by an authoring committee of experts. Reports typically include findings, conclusions, and recommendations based on information gathered by the committee and the committee's deliberations. Each report has been subjected to a rigorous and independent peer-review process and it represents the position of the National Academies on the statement of task.

Proceedings published by the National Academies of Sciences, Engineering, and Medicine chronicle the presentations and discussions at a workshop, symposium, or other event convened by the National Academies. The statements and opinions contained in proceedings are those of the participants and are not endorsed by other participants, the planning committee, or the National Academies.

Rapid Expert Consultations published by the National Academies of Sciences, Engineering, and Medicine are authored by subject-matter experts on narrowly focused topics that can be supported by a body of evidence. The discussions contained in rapid expert consultations are considered those of the authors and do not contain policy recommendations. Rapid expert consultations are reviewed by the institution before release.

For information about other products and activities of the National Academies, please visit www.nationalacademies.org/about/whatwedo.

Copyright National Academy of Sciences. All rights reserved.

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

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*

* Denotes member of panel's designated data analysis subgroup.

Copyright National Academy of Sciences. All rights reserved.

COMMITTEE ON NATIONAL STATISTICS

Katharine G. Abraham (*Chair*), University of Maryland, College Park
Mick P. Couper, Institute for Social Research, University of Michigan
Diana Farrell, JPMorgan Chase Institute, Washington, D.C.
Robert M. Goerge, Chapin Hall at the University of Chicago
Erica L. Groshen, School of Industrial and Labor Relations, Cornell University
Daniel E. Ho, Stanford Law School and Stanford Institute for Economic Policy Research
Hilary Hoynes, Goldman School of Public Policy and Department of Economics, University of California, Berkeley
Daniel Kifer, Department of Computer Science, Pennsylvania State University
Sharon Lohr, School of Mathematical and Statistical Sciences, Arizona State University, *Emeritus*
Nela Richardson, ADP Research Institute
C. Matthew Snipp, School of the Humanities and Sciences, Stanford University
Elizabeth A. Stuart, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health

Staff
Melissa Chiu, *Director*
Constance F. Citro, *Senior Scholar*
Brian Harris-Kojetin, *Senior Scholar*

Copyright National Academy of Sciences. All rights reserved.

Reviewers

This Consensus Study Report was reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise. The purpose of this independent review is to provide candid and critical comments that will assist the National Academies of Sciences, Engineering, and Medicine in making each published report as sound as possible and to ensure that it meets the institutional standards for quality, objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

We thank the following individuals for their review of this report:

RUSSEL E. CAFLISCH, Courant Institute of Mathematical Sciences, New York University
BARBARA ENTWISLE, Carolina Population Center, University of North Carolina at Chapel Hill
ROBERT E. FAY, Statistical Fellow, Westat, Inc.
IVAN P. FELLEGI, retired, Statistics Canada
RUOBIN GONG, Department of Statistics, Rutgers University
JERI GREEN, retired U.S. Census Bureau and Independent consultant, Washington, DC
CRISTOPHER MOORE, Resident Fellow, Santa Fe Institute
GREGG ROBINSON, retired U.S. Census Bureau and Independent consultant, Alexandria, VA
ALEXIS R. SANTOS, Department of Human Development and Family Studies, Pennsylvania State University
SAMUEL V. SCARPINO, Artificial Intelligence and Life Sciences and Practice, Health and Computer Sciences, Santa Fe Institute

Although the reviewers listed above provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recom-

Copyright National Academy of Sciences. All rights reserved.

mendations of this report nor did they see the final draft before its release. The review of this report was overseen by **SUSAN HANSON**, Department of Geography, Clark University, and **SUSAN A. MURPHY**, Department of Statistics, Harvard University. They were responsible for making certain that an independent examination of this report was carried out in accordance with the standards of the National Academies and that all review comments were carefully considered. Responsibility for the final content rests entirely with the authoring committee and the National Academies.

Copyright National Academy of Sciences. All rights reserved.

Disclosure Review Statement

The U.S. Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed the information products used to produce several figures and tables in this report for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to these releases.

By mutual agreement, Disclosure Review Board clearance was sought and obtained only for the generation of specific figures and tables published in this report, and not for the issuance of the underlying data files as standalone products or as addenda to this report.

For brevity, the SOURCE statement for each relevant figure or table in the text refers to this Disclosure Review Statement and lists the associated Disclosure Review Board approval number. These approval numbers are:

- CBDRB-FY23-0171 (Nonresponse Followup [NRFU] resolutions by phase, figures)
- CBDRB-FY23-0179 (race/ethnicity analysis)
- CBDRB-FY23-0180 (age heaping analysis)
- CBDRB-FY23-0197 (Master Address File [MAF] development, figures)
- CBDRB-FY23-0206 (correlation coefficients)
- CBDRB-FY23-0214 (NRFU resolutions by phase, tables)
- CBDRB-FY23-0215 (MAF development, tables)
- CBDRB-FY23-0221 (Group Quarters data quality)
- CBDRB-FY23-0224 (NRFU and Self-Response return rate analysis by ventiles of American Community Survey variables)

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

Acknowledgments

The Panel to Evaluate the Quality of the 2020 Census wishes to thank the many people who have contributed to its work and to this final report.

We appreciate the support of the U.S. Census Bureau and its encouragement of independent, external review of the 2020 Census, under the acting directorship of deputy director Ron Jarmin in 2021 and then under director Robert Santos since his assumption of the office in January 2022. Karen Deaver has been an excellent collaborator as the primary information conduit between the panel and the Census Bureau, and the panel has benefited from the active engagement and interaction of Albert Fontenot and Deborah Stempowski as heads of the decennial census directorate. Kristee Camilletti provided great assistance in overseeing the contract behind this study and Kawaldeep Kaur Brar was instrumental in guiding panel members and staff through the clearance process. In the late stages of the work, RJ Marquette and Kelly Mathews were vital in ushering our requests to use the results of data analyses behind the Census Bureau information technology firewall through the Disclosure Review Board process. We have also benefited greatly from interactions with key Census Bureau staff including John Abowd, Dominic Beamer, Dierdre Dalpiaz Bishop, Maryann Chapin, Stuart Irby, Sallie Keller, Rafael Morales, Jennifer Ortman, and Victoria Velkoff.

In addition to those already named, we appreciate the contributions of Census Bureau staff who presented and discussed material with the panel during its open meetings and in regular sessions with the panel's data analysis subgroup, including: Tamara Adams, Willette Allen, Robin Bachman, Judy Belton, Christine Flanagan Borman, Patrick Cantwell, Richard Denby, Jason Devine, Daniel Donello, Dora Durante, Monique Eleby, Christine Hartley, Michael Hawes, Kristen Hearns, Joan Hill, Cynthia Davis Hollingsworth, Eric Jensen,

Copyright National Academy of Sciences. All rights reserved.

Andrea Johnson, Nicholas Jones, Timothy Kennel, Sarah Konya, Julia Lopez, Barbara LoPresti, Frank McPhillips, Rachel Marks, Leanna Mayo, Thomas Mule, Roberto Ramirez, Jennifer Reichert, Rita Schuler, Kathleen Styles, Ben Taylor, Kevin Zajac, and Mary Frances Zelenak.

We reiterate our deep appreciation to the 2020 Census field staff and external census stakeholders who participated in the panel's public meetings in 2021 and refer readers to the acknowledgments and compilation of meeting agendas in our interim report for the full credits.

This report contains many original analyses that would not have been possible without sustained work by our analytic subgroup. Its members were Robert M. Bell, Thomas A. Louis, Joseph J. Salvo, and David Van Riper with subgroup chair Lloyd B. Potter. The analytic subgroup met virtually twice a week for nearly a full year, working behind the Census Bureau's information technology firewall and then preparing proposals to the Census Bureau to clear some of these analyses for publication. After this work continued outside the firewall, Margo Anderson and Marc Hamel joined most of the biweekly meetings. Wendy Underhill also attended many of these meetings.

The COVID-19 pandemic affected not only the careful planning of the Census Bureau, but the timeline and working conditions for this panel as well. We could not have done this work without a helpful and dedicated staff. Many members of the Committee on National Statistics (CNSTAT) and National Academies of Sciences, Engineering, and Medicine staff facilitated our work. Constance Citro, senior scholar, offered us wise counsel at every point and participated in all our sessions. Her knowledge and experience with the census proved invaluable to us. In particular, Connie's expert drafts of core chapters on coverage measurement, race and ethnicity measurement, and age heaping were vital in sharpening our thinking and advancing our work. Michael Cohen's assistance to the data analysis subgroup was critical to the group's progress, and he was also instrumental to our review of census operations and administrative data. Katrina Baum Stone brought her experience with several federal statistical agencies to bear on our work on the census and helped us find the forest when there were too many trees. Anthony Mann handled panel logistics, not a simple task given the need for all meetings to be virtual or hybrid, with various members and interviewees experiencing connectivity issues and time zone constraints. Brian Harris-Kojetin, CNSTAT director through February 2023, has served as the lead contact on the Census Bureau contract and has offered us helpful information at timely moments, and CNSTAT support of this panel and its work has continued under the directorship of Melissa Chiu.

Our heartiest thank-you goes to Daniel Cork, the study director, who dealt with the minutiae of multiple subtopics, shifting data points, and often conflicting suggestions from the panelists and others. With a report of many hundreds of pages, he managed the herculean effort of tracking chapter versions, coordinating the presentation of figures, and keeping the panel timely and on

Copyright National Academy of Sciences. All rights reserved.

budget. We appreciated his considerable analytic talent and his skill in making complex topics more understandable. His candor, objectivity, and perseverance made this panel's work possible.

Teresa A. Sullivan, *Chair*
Panel to Evaluate the Quality of the 2020 Census

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

Contents

Disclosure Review Statement **ix**

Abstract **1**

Summary **5**

1 Introduction **21**
1.1 The Census as Essential Civic Ceremony 21
1.2 Nature and Limitations of This Study 23
1.2.1 Panel Charge and Data Analysis Subgroup Structure 23
1.2.2 Defining Census Quality 24
1.2.3 The Enduring 2020 Census 26
1.2.4 Limits on Analysis and Publication 26
1.3 Overview of This Report 28

2 Overview of the 2020 Census **31**
2.1 Four Key Innovation Areas, and Initial Vision for the 2020 Census 31
2.2 Emergence of External Pressures 33
2.3 Evolution of the 2020 Census Design 35
2.3.1 Progression of 2020 Census Operational Plans 35
2.3.2 Immediate Precensus Challenges and Operational Revisions 40
2.4 The 2020 Census, Planned and Unplanned 42
2.4.1 The COVID-19 Pandemic and Resulting Schedule Changes 42

Copyright National Academy of Sciences. All rights reserved.

2.4.2 Presidential Memorandum on Undocumented Immigrant Counts, Schedule "Replan," and End of Data Collection 42
2.4.3 Changes in 2020 Census Operations Due to the COVID-19 Pandemic 43
2.4.4 2020 Census Data Products 47
2.5 Conclusion 48

3 Age Heaping as an Indicator of Data Quality 49
3.1 Age Heaping as a Quality Indicator for the 2020 Census 49
3.2 Analysis of Age Reporting in the Census 51
3.3 Sources of Age Heaping—Type of Response 55
3.4 Age Heaping by Response Mode 57
3.5 Age Heaping—Conclusion 61
3.6 Data Quality in the Nonresponse Followup Operation 64

4 Coverage Measurement in the 2020 Census 69
4.1 Estimating Coverage Errors 70
4.1.1 Demographic Analysis 70
4.1.2 Post-Enumeration Survey with Dual System Estimation 72
4.2 2020 Coverage Results 74
4.2.1 Total Population 74
4.2.2 Age and Sex 76
4.2.3 Race, Ethnicity, Housing Tenure, Age and Sex Groups, States 77
4.2.4 Post-Enumeration Survey—Components of Error 80
4.3 Cautions and Areas for Research 86
4.3.1 Uncertainty in Demographic Analysis 86
4.3.2 Delays in Demographic Analysis Net Undercount Estimates by Race 90
4.3.3 Lack of Subnational Demographic Analysis Estimates 91
4.3.4 Uncertainty in Post-Enumeration Survey Results 91
4.3.5 Comparison of Errors in the 2010 and 2020 PES 92
4.3.6 Limited Analysis of Components of Coverage Error 92
4.3.7 Limited Granularity of Coverage Estimates 95
4.4 Conclusions and Recommendations 96

5 Master Address File 101
5.1 Overview of Master Address File Development and the Dataset 103
5.2 MAF1 and MAF2, Precensus and Postcensus Pivot Points in the Master Address File 105
5.3 Provenance of Addresses in the 2020 Census Master Address File and Relation to Address-Building Operations 108

Copyright National Academy of Sciences. All rights reserved.

5.4 Spatial Variation in Master Address File Unit Categories 114
5.4.1 Legacy Sources 114
5.4.2 Good Reactivations 116
5.4.3 Unique Adds from 2020 Census Address-Building Operations 118
5.5 Impact of 2020 Census Address-Building Operations 121
5.5.1 Address List Input from User Programs 126
5.5.2 Field/Clerical Operations 129
5.5.3 Census Enumeration Operations 133
5.6 Conclusions 133

6 Self-Response to the Census 145
6.1 Self-Response Strategies and Processes 149
6.1.1 2010 Census Self-Response Procedures 149
6.1.2 2020 Census Self-Response Procedures 150
6.2 Analysis of Variation in Self-Response 153
6.2.1 Previous Work on Self-Response for Census Tracts 153
6.2.2 Panel's Analysis of Self-Return Rates 154
6.3 Conclusions and Recommendations 167

7 Nonresponse Followup 171
7.1 Overview of 2020 Nonresponse Followup Operation and Challenges 171
7.2 Descriptive Analysis of Nonresponse Followup Enumeration by Modality and Phase 173
7.3 Cluster Analysis of Variation in Nonresponse Followup Resolutions by Phase 175
7.4 Nonresponse Followup Enumeration Modalities and Demographic, Economic, and Housing Characteristics 180
7.4.1 Nonresponse Followup Correlations with Social and Economic Characteristics of Census Tracts 184
7.4.2 Nonresponse Followup Ventile Plots 184
7.5 Conclusions 195

8 Use of Administrative Records for Enumeration in the 2020 Census 201
8.1 Administrative Records Enumerations in the 2020 Census 203
8.2 Models Developed for This Purpose 204
8.2.1 Housing Unit Status Model 205
8.2.2 Models to Estimate the Number of Occupants 206
8.2.3 Person-Place Model 207
8.2.4 Household Composition Model 208
8.3 Joining the Three Models—Distance Functions 209

Copyright National Academy of Sciences. All rights reserved.

8.4 Evaluating the Approach Used in 2020 211
8.5 What the Evaluation Does Not Answer 214
8.6 Looking to 2030 216
8.6.1 Administrative Records Enumeration as Part of Nonresponse Followup 216
8.6.2 Using Administrative Records on a Large Scale for Census Enumerations—Examples from Peer Countries 219
8.6.3 Determining the Appropriateness and Feasibility of a Significantly Increased Role for Administrative Records Enumerations in the 2030 Census 220
8.6.4 "Real-Time" Simulation of an Administrative Records Census 223
8.7 Conclusions and Recommendations 225

9 Measurement of the Group Quarters Population 227
9.1 What Are Group Quarters and Why Are They Important? 227
9.2 Accurate Counts of Group Quarters Populations Matter for Local Governments 228
9.3 Number of Persons in Group Quarters by Type 229
9.4 Group Quarters Address List Development 232
9.5 Nature and Origin of Group Quarters Addresses 234
9.6 Postcensus Review 235
9.7 The Collection of Group Quarters Data in the 2020 Census 237
9.7.1 Occupancy Status 237
9.7.2 Imputation of Occupancy Status and Population Counts 238
9.7.3 Methods of Data Collection by Group Quarters Type 239
9.7.4 Facility Self-Enumeration 248
9.7.5 Item Nonresponse by Group Quarters Type 248
9.8 Conclusions and Recommendations 255

10 Measurement of Race and Ethnicity 261
10.1 Overview of Race and Ethnicity Measurement 262
10.1.1 U.S. Office of Management and Budget Standards and Census Race and Ethnicity Questions 262
10.1.2 Race and Ethnicity Distributions from 1970–2020—A More Diverse Nation 268
10.2 New Write-Ins, Data Capture, and Coding for 2020 272
10.2.1 Effects on Race and Ethnicity Estimates from the Census and the American Community Survey 272
10.2.2 Possible Causes 273
10.3 Coverage Error 276
10.4 Imputations for Missing and Inconsistent Data 277
10.5 2020 Disclosure Avoidance System (DAS) 278

Copyright National Academy of Sciences. All rights reserved.

10.5.1 Delays 279
10.5.2 Effects of Delays on Race and Ethnicity for Demographic Analysis and Population Estimates 280
10.5.3 Noise in the 2020 Redistricting File 283
10.6 Race and Ethnicity Measurement in 2020—Conclusions 284
10.6.1 Improvements to Collection of Race and Ethnicity Information 285

11 Impact of New Confidentiality-Protection Methods on 2020 Census Data Products 291
11.1 Census Data Products: Overview of Content and Timing of Release 295
11.2 Confidentiality Protection in the Decennial Census 300
11.3 The Road to the 2020 Disclosure Avoidance System 302
11.3.1 What Did the Operational Plan and Other Statements Say About Disclosure Avoidance? 302
11.3.2 Implementation Issues 305
11.3.3 Implications for Development of the 2020 Disclosure Avoidance System 311
11.4 Assessment 315
11.4.1 Fundamental Responsibilities of Statistical Agencies 316
11.4.2 Census Planning, Testing, and Innovation 321
11.4.3 Privacy Protection in Other National Statistical Offices 323
11.4.4 Effects on Other Census Bureau Programs 324
11.4.5 Effects on Operational Quality Metrics 325
11.5 Conclusions 327
11.6 Looking Ahead to 2030 329
11.6.1 Framework for Balancing Utility and Confidentiality 329
11.6.2 Data Product Plan for 2030 332

12 Learning from 2020, Preparing for 2030 335
12.1 Priority Goals for 2030 Census Research and Development 336
12.1.1 Maximize Census Self-Response 337
12.1.2 Improve Quality of Nonresponse Followup Data 339
12.1.3 Reduce Quality and Coverage Gaps by Race, Ethnicity, and Socioeconomic Status 340
12.1.4 Continuous Modernization of Geographic Resources for All Living Quarters 343
12.1.5 Realign Balance between Utility, Timeliness, and Confidentiality Protection in 2030 Census Data Products 344
12.2 Improving Census Research and Development 348
12.2.1 Continue to Learn from the 2020 Census Experience 348
12.2.2 Improving Census Experimentation and Testing 350

Copyright National Academy of Sciences. All rights reserved.

12.2.3 Systems Design to Facilitate Evaluation 352
12.3 Improving Census Processes 353
12.3.1 Continued Work on Administrative Records to Supplement Census Operations 353
12.3.2 Continued Development of Master Address File Curation and Updating 354
12.3.3 Improvements to Major Census Coverage Evaluation Methods 354
12.3.4 Revisit Unplanned Contingency Changes from 2020 that Merit Further Consideration 356
12.4 Improving Census Law 357

References 361

Appendixes 381

A Full List of Conclusions and Recommendations 383

B Glossary and Abbreviations 403

C Additional Detail and Reference on 2020 Census Operations 419

D Extensions of Census Coverage Evaluations 459

E 2020 Census Group Quarters Definitions and Type Code List 469

F Additional Detail and Reference on Data Products and Disclosure Avoidance 475

G Biographical Sketches of Panel Members and Staff 493

Copyright National Academy of Sciences. All rights reserved.

Boxes, Figures, and Tables

BOXES

1.1 Federal Committee on Statistical Methodology (2020) Data Quality Framework 25

2.1 Key 2020 Census Litigation 36
2.2 Type of Enumeration Areas and Mail Contact Strategies in the 2020 Census 38

3.1 Derivation of an Age Heaping Metric 53
3.2 Response Categories for Age and Date of Birth, 2010 and 2020 Censuses 56

4.1 Methods for Estimating Census Coverage 71
4.2 Major Steps in Post-Enumeration Survey Data Collection and Estimation 75

6.1 Self-Responses Higher Quality than Enumerator Responses 146
6.2 Primus for Internet Response 151

9.1 Group Quarters in Pennsylvania Townships and Counties 229
9.2 Post-Census Group Quarters Review Program 237
9.3 Imputation of Group Quarters Characteristics 249

10.1 Request for Comment on Proposed Race and Ethnicity Measurement Standards 267
10.2 Content of the 2020 Redistricting Data File 279

11.1 Methods for Protecting Confidentiality, 1970–2020 Census Summary Files 293
11.2 Effects of the 2020 Disclosure Avoidance System on Data Quality: Overview 294
11.3 Legislation Protecting the Confidentiality of Decennial Census Information 301
11.4 Basic Steps in TopDown Algorithm for 2020 Census Disclosure Avoidance 307
11.5 Statements of Statistical Agency Responsibilities 317

Copyright National Academy of Sciences. All rights reserved.

11.6 Types and Examples of Informational Materials Developed by the U.S. Census Bureau for the 2020 Disclosure Avoidance System 320
11.7 Relevant Language—Foundations for Evidence-Based Policymaking Act of 2018 330
11.8 Excerpts—Year 2 Report, Advisory Committee on Data for Evidence Building 331
11.9 Legislation on Shared Statistical Agency–Data User Responsibility for Privacy Protection 331

12.1 Major Changes to U.S. Code Title 13 and Census Law Since Codification of Title 13 in 1954 358

FIGURES

2.1 Simplified 2020 Census type of enumeration area and mailing strategy map, by county 39

3.1 Demographic Analysis estimates of percentage net under/overcount, by single years of age, 2010, 2020, middle series 51
3.2 Age heaping in the 2010 and 2020 Censuses, using the $\hat{h}_a$ metric 52
3.3 Age heaping for respondents with reported age only, 2010 and 2020 Censuses, using the $\hat{h}_a$ metric 57
3.4 Age heaping, all respondents except those with reported age only, 2010 and 2020 Censuses, using the $\hat{h}_a$ metric 58
3.5 Age heaping, responses based on whole-household imputations in 2010 and 2020 Censuses, using the $\hat{h}_a$ metric 58
3.6 Age heaping for respondents who used primary Self-Response option, 2010 Census (paper) and 2020 Census (internet), using the $\hat{h}_a$ metric 60
3.7 Age heaping for respondents obtained by administrative records in the 2020 Census, using the $\hat{h}_a$ metric 60
3.8 Age heaping for respondents enumerated in Nonresponse Followup, 2010 and 2020 Censuses, using the $\hat{h}_a$ metric 61
3.9 Age heaping for Nonresponse Followup resolutions by household member interview, 2010 and 2020 Censuses, using the $\hat{h}_a$ metric 62
3.10 Age heaping for Nonresponse Followup resolutions by proxy, 2010 and 2020 Censuses, using the $\hat{h}_a$ metric 62
3.11 Item nonresponse rates for age/date of birth, ethnicity, and race, by type of response, 2010 and 2020 Censuses 66

4.1 Demographic Analysis estimates of percentage net undercount or overcount, by single years of age and sex, 2020, middle series 77
4.2 Post-Enumeration Survey (2010, 2020) and Demographic Analysis (2020, middle series) estimates of percentage net undercount or overcount, by age and sex groups 78

Copyright National Academy of Sciences. All rights reserved.

4.3 Post-Enumeration Survey (2010, 2020) estimates of percentage net undercount or overcount, by race and ethnicity 79
4.4 Post-Enumeration Survey (2010, 2020) estimates of percentage net undercount or overcount for demographic groups 81
4.5 Percentages of erroneous census enumerations, whole-person imputations, and Post-Enumeration Survey people omitted from the census for the household population by self-response rate deciles for census tracts, 2020 Post-Enumeration Survey 86
4.6 Demographic Analysis estimates of percentage net undercount or overcount, by single years of age, 2020, low, middle, high series 87
4.7 Demographic Analysis estimates of percentage net undercount or overcount, by ethnicity, children and young adults, 2020, low, middle, high series 89

5.1 Basic formulation of MAF1 (entering 2020 Census enumeration) and MAF2 (2020 Census tabulation) filters 106
5.2 Percent of Master Address File units in Good Legacy Census category, by county 115
5.3 Percent of Master Address File units in Good Legacy Census category, by tract, Iowa 116
5.4 Percent of Master Address File units in Good Legacy Delivery Sequence File category, by county 117
5.5 Percent of Master Address File units in Good Legacy Delivery Sequence File category, by tract, Austin, Texas metropolitan area 118
5.6 Percent of Master Address File units in Bad Legacy category, by county 119
5.7 Percent of Master Address File units in Good Reactivations category, by county 120
5.8 Percent of Master Address File units in Good Reactivations category, by tract, Dallas–Fort Worth metropolitan area 121
5.9 Percent of Master Address File units in Good 2020 Adds category, by county 122
5.10 Percent of Master Address File units in Good 2020 Adds category, by tract, Los Angeles–Orange–San Diego Counties, California 123
5.11 Percent of Master Address File units in Bad 2020 Adds category, by county 124
5.12 Percent of Master Address File units in Bad 2020 Adds category, by tract, Los Angeles–Orange–San Diego Counties, California 125
5.13 Percent of Master Address File units added or touched by 2020 Census Local Update of Census Addresses operation, by county 128
5.14 Percent of Master Address File units added or touched by 2020 Census Local Update of Census Addresses operation, by tract, Baltimore–Washington metropolitan area 130
5.15 Percent of Master Address File units added or touched by 2020 Census Local Update of Census Addresses operation, by tract, Tampa–St. Petersburg metropolitan area 131
5.16 Percent of Master Address File units added or touched by 2020 Census Local Update of Census Addresses operation, by tract, Miami-Dade County metropolitan area 132

Copyright National Academy of Sciences. All rights reserved.

5.17 Percent of Master Address File units added or touched by 2020 Census Office-Based Address Verification operation, by county 134
5.18 Percent of Master Address File units added or touched by 2020 Census Office-Based Address Verification operation, by tract, Los Angeles–Orange–San Diego Counties, California 135
5.19 Percent of Self-Response housing units (occupied and vacant) using Non-ID Processing, by county 136
5.20 Percent of Self-Response housing units (occupied and vacant) using Non-ID Processing, by tract, Los Angeles–Orange–San Diego Counties, California 137
5.21 Percent of Master Address File units added or touched by 2020 Census Nonresponse Followup operation, by county 138
5.22 Percent of Master Address File units added or touched by 2020 Census Nonresponse Followup operation, by tract, Missouri 139
5.23 Percent of Master Address File units added or touched by 2020 Census Update Leave operation, by county 140
5.24 Percent of Master Address File units added or touched by 2020 Census Update Leave operation, by tract, San Bernardino–Riverside–San Diego–Imperial Counties, California 141

6.1 Percentages of people omitted from the census for the household population by self-return rate deciles for census tracts, 2000 and 2020 Census Post-Enumeration Surveys 147
6.2 2020 Census percent of item nonresponse for age/date of birth, ethnicity, and race by Self-Response and Nonresponse Followup (NRFU) modalities 148
6.3 Zero-order correlations between total Self-Response return rate and American Community Survey socioeconomic characteristics, census tracts 156
6.4 Zero-order correlations between Internet Self-Response return rate and American Community Survey socioeconomic characteristics, census tracts 157
6.5 Zero-order correlations between Paper Self-Response return rate and American Community Survey socioeconomic characteristics, census tracts 158
6.6 Zero-order correlations between Telephone Self-Response return rate and American Community Survey socioeconomic characteristics, census tracts 159
6.7 Mean Self-Response return rate by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) average household size and (b) percent bachelor's degree or higher 161
6.8 Mean Self-Response return rate by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) median household income and (b) unemployment rate 162
6.9 Mean Self-Response return rate by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent non-Hispanic White alone and (b) percent non-Hispanic Black or African American alone 163
6.10 Mean Self-Response return rate by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent Hispanic or Latino (any race) and (b) percent speak English less than very well 164

Copyright National Academy of Sciences. All rights reserved.

6.11 Mean Self-Response return rate by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent housing units renter occupied and (b) percent vacant housing units 165
6.12 Mean Self-Response return rate by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent structures with 20 or more units and (b) percent with broadband internet subscription 166

7.1 Nonresponse Followup resolutions by household interview/observation, proxy, administrative records, and special means, by operational phase, 2020 Census Field Supervisor Areas 176
7.2 Percent of workload resolved in Phases 2–3 minus percent resolved in Phase 1, by major resolution mode and outcome, by county 177
7.3 Summary Nonresponse Followup resolutions by phase (combining Phases 2 and 3), mode, and outcome, by county 178
7.4 Summary Nonresponse Followup resolutions by phase (combining Phases 2 and 3), mode, and outcome, by tract 179
7.5 *K*-means clustering of county-level Nonresponse Followup resolution composition percentages into three clusters 181
7.6 Geographic distribution of *k*-means clustering of Nonresponse Followup resolution composition percentages, by county 182
7.7 Geographic distribution of *k*-means clustering of Nonresponse Followup resolution composition percentages, by tract 183
7.8 Zero-order correlations between percent Nonresponse Followup resolutions by household member interview and American Community Survey socioeconomic characteristics, census tracts 185
7.9 Zero-order correlations between percent Nonresponse Followup resolutions by administrative records and American Community Survey socioeconomic characteristics, census tracts 186
7.10 Zero-order correlations between percent Nonresponse Followup resolutions by proxy and American Community Survey socioeconomic characteristics, census tracts 187
7.11 Zero-order correlations between percent Nonresponse Followup resolutions as "population count only" and American Community Survey socioeconomic characteristics, census tracts 188
7.12 Mean percent Nonresponse Followup resolutions by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) average household size and (b) percent bachelor's degree or higher 189
7.13 Mean percent Nonresponse Followup resolutions by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) median household income and (b) unemployment rate 190
7.14 Mean percent Nonresponse Followup resolutions by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent non-Hispanic White alone and (b) percent non-Hispanic Black or African American alone 191

Copyright National Academy of Sciences. All rights reserved.

7.15 Mean percent Nonresponse Followup resolutions by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent Hispanic or Latino (any race) and (b) percent speak English less than very well 192
7.16 Mean percent Nonresponse Followup resolutions by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent housing units renter occupied and (b) percent vacant housing units 193
7.17 Mean percent Nonresponse Followup resolutions by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent structures with 20 or more units and (b) percent with broadband internet subscription 194

9.1 Person data collected using Facility Self-Enumeration by group quarters type and method of collection, 2020 Census 248
9.2 2010 and 2020 Census item nonresponse rate comparison of person-level items, group quarters enumerations 250
9.3 Item nonresponse rates by group quarters type, data collection via eResponse 252
9.4 Hispanic-origin item nonresponse rates by percent responding by eResponse, college/university student housing, 50 states and District of Columbia 254
9.5 Hispanic-origin item nonresponse rates by percent Hispanic, college/university student housing, 50 states and District of Columbia 255
9.6 Item nonresponse rates by group quarters type, data collection via Paper Listing 256

10.1 Percentage distribution of the population by race and ethnicity, 1970–2020 Censuses 269
10.2 Percentage Hispanic and percentage distribution by race for people over and under age 18, 2010 and 2020 Censuses 271
10.3 Percentage distribution of response modes for race and ethnic groups, 2020 Census 274
10.4 Ratio of write-in race responses (for people who provided at least one write-in) to checkbox-only responses by response mode, total and Hispanic population, 2020 Census 275
10.5 Percentage of counties, incorporated places, and census tracts with differences of 10% or more between the 2020 Disclosure Avoidance System privacy-protected and original 2010 redistricting files in estimates for race and Hispanic populations 284

11.1 Population at the 10th–90th percentiles, minor civil divisions, incorporated places, and American Indian reservations, 2020 Census 311
11.2 Hierarchy of census geographic entities 312

Copyright National Academy of Sciences. All rights reserved.

TABLES

2.1 Requested and Enacted Budget for the 2020 Census, by Fiscal Year, and Date of Enactment 34

3.1 Overall Age Heaping (Ages 23–62): Modified Whipple's Index (MWI) by Sex and Racial-Ethnic Group, 2010 and 2020 Censuses 55
3.2 Overall Age Heaping (Ages 23–62): Modified Whipple's Index (MWI), 2010 and 2020 Censuses, by Age/Date of Birth (DOB) Reporting Status 56
3.3 Overall Age Heaping (Ages 23–62): Modified Whipple's Index, 2010 and 2020 Censuses, by Response Mode 59
3.4 Percentage of the Household Population by Components of Coverage Error by Census Operation, 2020 Census Post-Enumeration Survey 65

4.1 Components of the Dual-System Estimator of the True Population 73
4.2 Percentage Net Coverage Error for the Total Resident Population: 2010 and 2020 Demographic Analysis and Post-Enumeration Survey Estimates 76
4.3 Net Coverage Rates and Differential Net Undercount Rates Comparing White Non-Hispanic People to Specified Groups, 2010 and 2020 Post-Enumeration Survey 80
4.4 Percentages of the Household Population by Components of Coverage Error by Race and Ethnicity, 2010 and 2020 Census PES 83
4.5 Percentage of the Household Population by Components of Coverage Error by Census Operation, 2020 Census Post-Enumeration Survey 85
4.6 Schedule of Person Interviewing Field Operations, 2010 and 2020 Post-Enumeration Surveys 93
4.7 2010 and 2020 Post-Enumeration Survey Features Relevant for Error 94

5.1 MAF1 (Entered Enumeration Process in 2020 Census) and MAF2 (Actual Enumeration in 2020 Census) Filters 107
5.2 Master Address File Entries by Typology Based on MAF1/MAF2/Collection Attempt Flags and Address Origin 110
5.3 Disposition of 2020 Census Collection Universe Master Address File Entries by Address Origin and Final 2020 Census Status 113
5.4 Addresses Added or Touched by 2020 Census Address-Building Operations, by Address Origin 127

7.1 National Summary of 2020 Census Nonresponse Followup Resolutions by Major Phase 174

9.1 Change in Group Quarters Population by Major Group Quarters Type, 2010 and 2020 Censuses, United States 231
9.2 Origins of Group Quarters Addresses in the 2020 Census 235
9.3 Transitions and Additions for Group Quarters Master Address File Units in the 2020 Census 236
9.4 Group Quarters by Occupancy Status and Type, 2020 Census 238
9.5 Count Imputation for Group Quarters, 2020 Census 240

Copyright National Academy of Sciences. All rights reserved.

9.6 Allowable Enumeration Methods by Group Quarters Type and Code, 2020 Census 244
9.7 Persons in Group Quarters by Group Quarters Type and Method of Data Collection, 2020 Census 246
9.8 2020 Census Item Nonresponse Rates for Group Quarters Person-Level Items, by Type 249
9.9 Hispanic-Origin Item Nonreponse Rates, eResponse for College/University Student Housing, by State 253

10.1 U.S. Office of Management and Budget Standard Race and Ethnicity Categories in Statistical Policy Directive No. 15 263
10.2 Response Categories to Race and Ethnicity Questions, 1970–2020 Censuses 264
10.3 Percentage Non-Hispanic/Hispanic of Total Population and Percentage Distribution of the Non-Hispanic/Hispanic Population by Race, 1970–2020 Censuses 270
10.4 Percentage Distribution of the Population by Race and Ethnicity, 2010–2020 Censuses and 2019–2020 1-Year American Community Survey Estimates 273
10.5 Percentage of People with at Least One Item Reported Who Had Ethnicity Imputed and Percentage Who Had Race Imputed, by Response Mode, 2010–2020 Censuses 277
10.6 Percentage Distribution of the Population by Race/Ethnicity, 2020 Population Estimates, 2020 Census 282
10.7 Percentage Distribution of the Population by Race/Ethnicity, 2010–2020 Censuses and Combined Race-Ethnicity Question 287

11.1 Release Dates/Scheduled Dates for Four Census Data Products, 1990–2020 Censuses 297
11.2 Features of the TopDown Algorithm Compared with the SafeTab-P, SafeTab-H, and PHSafe Algorithms 314

12.1 Response Rate (%) Via Online Forms in Recent International Censuses 338

C.1 Modification of Timing of 2020 Census Enumeration Operations Due to COVID-19 432
C.2 2020 Census Self-Response Mail Contact Strategy and Adjustments for COVID-19 434
C.3 2020 Census Non-ID Workload by Process and Outcome 437
C.4 Item Imputation Rates (Percentages), 2010 and 2020 Censuses 452
C.5 Differences in Numbers of Write-Ins for the Race and Hispanic-Origin Questions, 2010 and 2020 Censuses 454

D.1 Hypothetical Example of How a Specific 5-Person Household Would Appear in Various Data Sources 466

F.1 Census Data Products, Content, Geographies, and Release Schedule, 2010–2020 479

Copyright National Academy of Sciences. All rights reserved.

F.2 Demonstration Product Releases and Other Events in the Implementation of the 2020 Disclosure Avoidance System, 2019–2023 483

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

Abstract

The census is a cornerstone of American democracy because political representation is explicitly tied to population counts. Census data also underlie government and business planning, allocation of billions of federal dollars to states and localities, and countless public and private uses, such as forming the denominators for public health rates. Thus, the quality of a decennial census, in terms of fulfilling these fundamental purposes, is necessarily difficult to assess before all major data products from that census are developed and released. Yet such is the task of this Panel to Evaluate the Quality of the 2020 Census, in light of the truly extraordinary circumstances of the 2020 Census of the United States. In this report, we provide a detailed explication of a few fundamental messages, which we succinctly outline here.

The overriding, signature achievement of the 2020 Census is that there was a 2020 Census at all. This statement is meant as high praise for the work of U.S. Census Bureau staff in conducting the census amidst an unceasing array of challenges, notably a months-long delay in field operations due to the onset of the COVID-19 pandemic. The Census Bureau successfully carried out a census under exceptionally difficult circumstances not only because of the hard work of its staff, but also because it had honed its development efforts for 2020 on a small number of innovation areas. These key innovations were enabling self-response via the internet, shifting most precensus Address Canvassing work from field visits to in-office review of imagery and other data, reengineering field management and case handling systems, and permitting the use of administrative records data to provide enumerations for some nonresponding households. As the census approached, the Census Bureau made many scope adjustments to its planned operations, and the pandemic and other adverse conditions forced a great many other revisions. Collectively, the focus on a small number of innovations helped the Census Bureau react as well as could be expected to the difficult conditions that faced 2020 Census operations.

Copyright National Academy of Sciences. All rights reserved.

Error—in the statistical sense of deviation from the unknown true value—is an inevitable part of the census, and the complete absence of error is an unrealistic and unattainable standard. Accordingly, it is not surprising (nor is it meant as criticism of Census Bureau work) that there were distinct quality problems with the 2020 Census, especially when compared with the 2010 Census. The most vivid illustration of these problems is the phenomenon of age heaping in 2020 Census results (the disproportionate reporting of age in approximate, round-sounding values ending in 0 or 5), which our analyses demonstrate to be principally a function of proxy reporting in the Nonresponse Followup (NRFU) operation. The quality of the count of college and university students was particularly hard-hit by the timing of the COVID-19 pandemic and the fieldwork stoppage. Lastly, compared with 2010, increased differential coverage with respect to race and ethnicity (Hispanic origin) was present in the 2020 Census, based on analyses of 2020 Census self-return rates, NRFU operational data, and the 2020 Post-Enumeration Survey.

Regardless of unique challenges forced upon the 2020 Census, other challenges resulted from the Census Bureau's own decisions. Chiefly, in the very late stages of 2020 Census planning, the Census Bureau decided to completely replace its methods for protecting the confidentiality of census data (termed Disclosure Avoidance System [DAS]) with an entirely new approach—one that had not been tested, prototyped, or deployed in the population census context. While confidentiality protection is a critically important responsibility of a statistical agency, this decision was made without appropriate consideration and balance regarding the utility of resulting census data products to fulfill the many important functions of census data. In short, the new DAS was not ready for use in 2020 Census production and substantially degraded the value of 2020 Census data products in terms of both quality and timeliness. Moreover, lingering questions about both the simulated database reconstruction attack that motivated the new DAS and the degree of confidentiality protection that was ultimately realized through final parameter settings have arguably harmed 2020 Census data products, some of which are not set to be released until late 2024, and the reputation of the Census Bureau.

Envisioning the 2030 Census and beyond, we echo the general guidance of previous National Academies of Sciences, Engineering, and Medicine panels: to effect major change in census conduct, the Census Bureau should 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 them. As priority goals for 2030 research and development, we suggest: (1) maximize self-response to the census; (2) improve the quality of data in Nonresponse Followup, including reduction if not elimination of low-confidence proxy reporting when a good alternative is available; (3) reduce gaps in coverage and data quality associated with race, ethnicity, and socioeconomic status; (4) improve the quality of address listings and contact strategies for

Copyright National Academy of Sciences. All rights reserved.

conventional housing units and group quarters locations alike; and (5) realign the balance between utility, timeliness, and confidentiality protection in 2030 Census data products. In addition to the five priority goals, we encourage research and development of ways to extend the use of administrative records to support the 2030 Census along the lines of the 2020 Census. We do not, however, support major movement toward an administrative records-based census as part of the 2030 Census design. That movement, if desired, is a much longer-term proposition for which the research, evidentiary, and legal base must be carefully developed. In the panel's judgment, goals and designs for the 2030 Census should be developed in true partnership with census data users and the myriad community of stakeholders and state, local, tribal, and federal government partners that make the census the vital, grand civic ceremony that it is.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

Summary

The survival of a democracy necessarily requires that successive generations of the population embrace its institutions, procedures, and practices. Since 1790, the U.S. census has been such an institution. It is a reoccurring civic ceremony in which everyone counts, and it reaffirms a commitment to equality among all. The census is a cornerstone of American democracy because political representation is explicitly tied to population counts. The constitutional purpose of the census is apportioning the U.S. House of Representatives among the states. Congress enacted Public Law 94-171 [1975] to ensure that the U.S. Census Bureau provides relevant census data to the states for redistricting. The census also serves as a vehicle for fairness in other contexts, such as the distribution of funds for federal programs. Census data underlie government and business planning, as well as countless private and public uses such as forming the denominators for public health rates.

It is within this context of the importance and multifaceted purposes of the decennial census that the charge for assessing the quality of the 2020 Census was set forth in the Statement of Task for this consensus study. Specifically, the Census Bureau sponsored this study with a very general but expansive statement of task delineating three key requirements:

1. Review information from the Census Bureau on the data collected as well as various process measures and indicators of data quality obtained as part of 2020 Census operations;
2. Review other available information, such as results from Demographic Analysis (DA), process measures and preliminary results from the Post-Enumeration Survey, and analyses of administrative records; and
3. Consider the results from evaluations of similar indicators from the 2010 and 2000 Censuses.

In this final report of the Panel to Evaluate the 2020 Census, we build on the frameworks for understanding census quality developed in our interim report,

Copyright National Academy of Sciences. All rights reserved.

bringing to bear the results of original data analysis performed by a dedicated subgroup of the panel. This panel was granted unprecedented access to data and resources, with the subgroup performing its work behind the Census Bureau's information technology firewall; we are grateful for this access and acknowledge that very few if any of the data and information requests we made were denied. That said, we also note that all presented results and analyses from our work have been subject to the Census Bureau's standards for information disclosure review. Hence, we have endeavored to be as thoroughgoing as possible in our analyses, but are necessarily limited in what we can report, constrained not only by the lack of available time and resources to address all areas of interest, but also by the mechanics and constraints on reported data of the Census Bureau's disclosure avoidance policies and infrastructure.

It is also important to note at the outset that it is inherently difficult, if not impossible, to fully evaluate a census while it is still in progress, yet that is the case with the 2020 Census. The 2020 Census data products have been greatly delayed, principally (as we discuss in detail) due to complications in implementing new techniques for confidentiality protection, and the Census Bureau's own evaluations and assessments have been delayed as well. Accordingly, it is inevitable that there will be discrepancies between what we say in this report and what is learned in the final 2020 Census data products and assessment reports.

The overriding, signature achievement of the 2020 Census is that there was a 2020 Census at all. This statement is meant as high praise for the work of Census Bureau staff in conducting the census amidst an unceasing array of challenges, notably a months-long delay in field operations due to the onset of the COVID-19 pandemic. The Census Bureau successfully carried out a census under exceptionally difficult circumstances, not only because of the hard work of its staff, but also because it had honed its development efforts for 2020 on a small number of innovation areas. These key innovations were enabling self-response via the internet, shifting most precensus Address Canvassing work from field visits to in-office review of imagery and other data, reengineering field management and case handling systems, and permitting the use of administrative records data to provide enumerations for some nonresponding households. In this report, we provide detailed chapters on these innovation areas and other major aspects of the 2020 Census in building a comprehensive assessment—starting with examination of a longstanding indicator of census and survey quality. This Summary describes each in turn, stating the full text of some (but not all, for brevity) of our conclusions and recommendations.[1]

[1] A full list of conclusions and recommendations is presented in Appendix A, and a Glossary and List of Abbreviations to aid the reader follows in Appendix B.

Copyright National Academy of Sciences. All rights reserved.

AGE HEAPING AS AN INDICATOR OF DATA QUALITY

Age heaping refers to a phenomenon whereby the ages of some survey subjects are misreported as favored ages—typically, but not necessarily, those ending in 0 or 5. Age heaping is important for two reasons. First, some uses of census data call for detailed age distributions at the state and local level that are not available from DA. Second, age heaping may be an indicator of other additional data-quality issues in the census, affecting both coverage and characteristics.

> *Conclusion 3.1: Age heaping (excess responses of ages ending in 0 or 5) is a well-recognized indicator of data quality in censuses. Age heaping was much more prevalent in the 2020 Census than in 2010, by an estimated factor of at least 2.5 for ages between 23 and 62 years, for which heaping is most common. This increase was almost completely due to increases in the rate of heaping during Nonresponse Followup (NRFU)—especially from proxy respondents but also from NRFU household respondents. Age heaping was rare in both censuses for self-responses, whether by paper or internet (in 2020 only), and there was no evidence of age heaping associated with administrative records.*

> *Conclusion 3.2: Findings about age heaping raise concerns about the overall quality of data collected during 2020 Nonresponse Followup (NRFU), especially from proxy respondents but even from some household members. Other quality indicators, including item nonresponse rates, numbers of people requiring imputation of all characteristics, and coverage errors, reinforce the difficulties of obtaining accurate data in NRFU and indicate that the difficulties increased in 2020 compared with 2010. The greater extent of quality problems in NRFU in 2020 compared with 2010 may be attributable to the COVID-19 pandemic (including the resulting delay in the NRFU schedule), but other factors might also be involved, which would need to be addressed in 2030 Census planning.*

COVERAGE MEASUREMENT IN THE 2020 CENSUS

Every census since 1790 has missed some people who should have been counted, counted some people twice, and counted other people who should not have been counted (e.g., someone born after Census Day). The balance of undercounted and overcounted people determines whether the census experienced a net undercount or overcount for the nation as a whole. In addition, people can be missed in one place and erroneously counted in another place according to census rules regarding usual residence. Also, people may misreport information such as age, or fail to report race or ethnicity, for

Copyright National Academy of Sciences. All rights reserved.

example, and have race or ethnicity erroneously imputed. These kinds of errors will not affect coverage for the nation but will add to net under- or overcount in a geographic area or population group. "Differential" net undercount or overcount—that is, the difference in coverage rates between population groups or geographic areas—is more concerning than the national net undercount or overcount. Principal methods for estimating census coverage (Box 4.1) include DA and the Post-Enumeration Survey (PES).

Conclusion 4.2: Comparing the 2010 and 2020 Census population coverage estimates from the Post-Enumeration Survey (PES):

- *Differences among race and ethnic groups widened substantially in 2020, with adverse implications for use of the data to allocate fixed resources, such as representation, funding, and services. Some groups saw increases in net overcounts, particularly non-Hispanic White Alone people and Asian people, and other groups saw increases in net undercounts, particularly Black people, Hispanic people, and American Indians and Alaska Natives on reservations.*
- *Increases in net undercounts in 2020 for groups such as Black and Hispanic people were driven by increases in omissions from the census (as measured by a lower match rate for people in the P-sample to the E-sample in the PES). Omissions, in turn, were higher in census tracts with lower rates of self-response.*
- *Net overcounts and undercounts were more pronounced for people of various groups in rental compared with owner housing, particularly in 2020 compared with 2010.*

Recommendation 4.1: For the Demographic Analysis (DA) coverage evaluation method, the U.S. Census Bureau should conduct research on:

- **The reasonableness of the assumption for incompleteness of Medicare enrollment in the high series, which produced coverage estimates that diverged greatly from the low and middle series for people ages 75 and older, through an appropriate match study;**
- **The suitability of Medicare data for coverage evaluation of the entire population ages 65 and older as in the 2010 Census and for improving the census count of this age group in 2030;**
- **The reasonableness of the methods for assigning Hispanic and non-Hispanic births for coverage estimates for children and young adults, which produced wide differences between the low and high DA series in 2020;**

Copyright National Academy of Sciences. All rights reserved.

- **Methods for narrowing estimates of net international migration, which affect estimates for Hispanic people and the total population; and**
- **Methods for developing experimental subnational DA estimates, starting with young children.**

Recommendation 4.3: For the Post-Enumeration Survey (PES) method, the U.S. Census Bureau should:

- **Perform sensitivity analyses, based on plausible assumptions, to put error bounds around such key operations as matching and imputation of match and enumeration status, to evaluate the quality of the 2020 PES and plan such analyses from the outset for the 2030 PES;**
- **Plan analyses for the 2030 PES of components of error for census operations in exhaustive and mutually exclusive categories that are as comparable as possible between 2020 and 2030;**
- **Seek adequate funding to increase the PES sample size in 2030 to at least 2010 levels; and**
- **Experiment with modeling to estimate net undercoverage and overcoverage for more detailed geographic and demographic groups for which direct estimates could not be provided in 2020.**

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, it follows that the quality of the decennial census is inextricably linked to the quality of the Census Bureau's geographic resources and, in particular, the address list that is used to mount census data-collection efforts. Prior to the 2000 Census, the address list used as the operational frame to conduct and monitor progress in the decennial census was constructed anew each decade, prior to and during the census. 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 intended to be a repository for all living quarters in the United States, including not only housing units but also group quarters (such as college and university student housing, health care/nursing facilities, and correctional facilities) and transitory locations (such as marinas, recreational vehicles, and hotels/motels). The most striking impression from analysis of metadata for the 2020 MAF is also the least surprising: the MAF is largely composed of addresses that

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

have been on the rolls for many years, and continuity in the MAF over time necessarily dominates change and churn.

Conclusion 5.1: The U.S. Census Bureau's approach of retaining a fluid Master Address File database and filtering operational extracts from that database is a sound one, provided that the filters are continually studied and updated to improve their efficacy.

Recommendation 5.1: The U.S. Census Bureau should continue to research and refine the filters it applies to the Master Address File to derive functional operational extracts, with the intent to reduce the number of addresses cancelled during collection (i.e., flagged as deletes in Nonresponse Followup and other field operations). Along those lines, the Census Bureau should continue to study ways to partner with and more fully utilize U.S. Postal Service data (including the Delivery Sequence File) and potential sources for address addition and revision information between censuses, such as address data recoverable from administrative records extracts, with particular eye toward more regular updating in areas without mail delivery.

Recommendation 5.2: The rules and scope of participant-input address-building operations of the census require clarification, including the Local Update of Census Addresses, New Construction, and Count Review. The U.S. Census Bureau should continue to improve ways to work with user-supplied input or data resources, and make it easier for state, local, and tribal authorities to supply input in usable form, while more clearly laying out the expectations for those address-building operations.

SELF-RESPONSE TO THE CENSUS

Two major census quality indicators are coverage (all eligible households are counted once and in the right place) and the quality of the data collected from each household in terms of accuracy and completeness. Evaluations have uniformly shown, including in 2020, that self-responses to the census are higher quality than those obtained in Nonresponse Followup (NRFU) and other in-person enumeration operations. The Census Bureau had a major success in 2020 with the use of the internet as a self-response mode in addition to paper and telephone and in achieving a high return rate (self-responses as a percentage of occupied housing units). The Census Bureau used a "wave" mailing strategy (multiple mailings) to encourage self-response—pushing

Copyright National Academy of Sciences. All rights reserved.

internet response (Internet First) in areas more equipped to use that response mode (e.g., areas with more households with broadband) and giving other areas the option up front to respond by paper or internet (Internet Choice). People could also respond via the internet without using the census ID number on printed materials (Non-ID enumerations).

Item nonresponse is an important data-quality element. Self-Response generally yields lower item nonresponse than Nonresponse Followup (NRFU) (see Figure 6.2). In comparing methods of Self-Response, the internet appears to have less item nonresponse for most data items. Telephone Self-Response appears to have slightly more item nonresponse compared with internet, though very few households used the telephone to respond. Paper has the most item nonresponse among the 2020 Self-Response modes, except for the age item.

> *Conclusion 6.1: Going from a paper-only 2010 Census to a majority Internet Self-Response 2020 Census is a commendable achievement, and it is particularly commendable that the U.S. Census Bureau's internally developed Primus system was able to be swapped in immediately prior to the census with excellent performance throughout the count. (Primus was the backup system pressed into primary service after concerns surfaced regarding the outsource-developed system.) The distinction between Internet First and Internet Choice contact strategies was a reasonable starting position.*

> **Recommendation 6.1: The U.S. Census Bureau should continue to investigate the performance of the contact strategies used in the 2020 Census, to increase overall self-response (particularly via the internet) in 2030, more in concordance with international experience with online census response. This investigation should include review of the effectiveness of the 2020 Census's segmentation into Internet First and Internet Choice strategies, as well as the timing and the operative contact/enumeration strategy (e.g., Internet First or Internet Choice mailings or Update Leave delivery) underlying successful Non-ID enumerations in 2020.**

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. There was also geographic variation in the proportion of households enumerated by each NRFU response mode. Overall, NRFU enumerations were of poorer quality than self-responses, with some NRFU response modes worse than others. Geographic variation in the quality of data

Copyright National Academy of Sciences. All rights reserved.

collected is likely associated with variations in populations' social, economic, demographic, and housing characteristics.

Conclusion 7.4: Nonresponse Followup (NRFU) enumerations result in more item nonresponse and poorer coverage than self-responses. NRFU item nonresponse for race and Hispanic origin exceeded 15% in the 2020 Census. Proxy enumerations had greater than 35% item nonresponse for race and Hispanic origin. Results suggest that census tracts with higher proportions of people with lower socioeconomic status (SES) characteristics were more likely to be enumerated using modes that resulted in higher item nonresponse. The net result is that 2020 Census data on race, Hispanic origin, and age are less accurate in those census tracts than in those with higher proportions of higher SES characteristics. This result is most relevant for redistricting. State and local officials are defining political districts with data that vary in how accurately they reflect the age and racial and ethnic composition of the population. Census data are particularly prone to error in areas with high percentages of people who are Black, Hispanic, or American Indian and Alaska Native living on reservations.

Recommendation 7.1: As a goal for the 2030 Census, the U.S. Census Bureau should consider a major reduction in the use of proxy interviewing for enumeration, if not the elimination of proxy reporting in all but very limited circumstances. Work toward this goal should be predicated on the results of research that only the Census Bureau can conduct, reexamining and comparing proxy reports in the 2020 Census with the information that would have been available from administrative records or third-party data resources.

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 Master Address File and in the Demographic Analysis method of coverage evaluation, but not to replace self-response or NRFU. In the 2020 Census, use of administrative records for enumeration in cases where a self-response was not received and one enumerator NRFU visit also failed to produce a return was a successful innovation. Administrative records generated almost 4% of occupied housing unit enumerations and over 14% of vacant housing unit enumerations. Further research and testing to determine how to expand this application of administrative records to reduce the NRFU workload is warranted. However, movement toward an administrative records-based census—enumerating a substantial proportion of

Copyright National Academy of Sciences. All rights reserved.

the population using records data without a contact attempt—is a long-term proposition that does not appear feasible for the 2030 Census.

Conclusion 8.1: Based on the information currently available, the U.S. Census Bureau deserves commendation for the meaningful steps made in the limited use of administrative records as an alternative to the enumeration of some nonresponding households in the 2020 Census. The Census Bureau acted with appropriate levels of both boldness and caution in settling on conditions under which a nonresponding household at a designated residence would be enumerated using records information after only one enumerator visit rather than six, yet still ensuring that single in-person visit. Moreover, this change in approach was preceded by an extensive history of high-quality experimentation and testing.

Conclusion 8.2: Moving to the use of administrative records to enumerate a substantial proportion of the population without a contact attempt does not seem feasible for the 2030 Census. Rather, it appears to be a long-term proposition, requiring extensive testing, research, and engagement with Congress, other stakeholders, and the public. Not only are there operational and technical issues to resolve and data-quality standards to determine and assess for geographic areas and population subgroups, but there is also the need to consider the ramifications of switching from a household/address-based focus to a person-based focus and to address issues of legal, constitutional, and public acceptability. It will be essential for the U.S. Census Bureau to be as transparent as possible about its research and testing and to have open and constructive dialogue with all parties.

Recommendation 8.1: The U.S. Census Bureau should continue its research and development program concerning the best ways to use administrative records as a supplement to decennial census operations. Potential uses of administrative records include expanding enumeration of limited subsets of the 2030 Census population in the Nonresponse Followup workload, reducing proxy responses and whole-person imputations, and possibly redressing the long-standing net undercount of children ages 0–4.

MEASUREMENT OF THE GROUP QUARTERS POPULATION

Group quarters (GQs) are places where people live or stay in a group living arrangement, and that are owned or managed by entities or organizations providing housing and/or services for the residents. These services may

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

include custodial or medical care as well as other types of assistance. GQs differ from typical household living arrangements because the people living in them are usually not related to one another. Moreover, GQs also include those who have no fixed address and are enumerated at places where they receive services, such as in shelters for those experiencing homelessness. Group Quarters Enumeration is always difficult, and the COVID-19 pandemic-induced operational adjustments were particularly hard to grapple with in this respect. As with other parts of the 2020 Census, Census Bureau staff rendered laudable and extraordinary efforts to adapt to the conditions and make the GQ count as complete as possible. Yet it is only fair to reiterate that not all of the 2020 Census challenges in Group Quarters Enumeration were unique to the circumstances of 2020.

Conclusion 9.1: The enumeration of nonhousehold group quarters (GQ) locations in the 2020 Census experienced major shortcomings, some of which can be directly attributed to COVID-19 pandemic adjustments to the data-collection scheduling (e.g., the scuttling of prearranged plans for in-person enumeration worked out in the Group Quarters Advance Visit operation). Yet several important operational and conceptual shortcomings cannot be attributed solely to the pandemic, such as the mismatch between census content and the U.S. Department of Education's interpretation of permissible data provision under federal privacy law, the shortfalls in the enumeration of people in military GQs, as well as the fluctuation in status (and resulting problems in the address list) in more fluid GQ types like group homes.

Recommendation 9.2: Group quarters (GQs) require more than updated address information; they require continually updated contact strategies as well. The U.S. Census Bureau should develop "selected advance contact" protocols with GQ facilities by type (and their parent companies, as appropriate), akin to follow-ups performed with businesses, and state, local, and tribal partners should be engaged in this update activity.

Recommendation 9.3: In light of the 2020 experience, the U.S. Census Bureau should convene group quarters (GQ) stakeholders in a discussion of more effective and useful options for electronic provision of data for GQ residents, and on the issues involved in obtaining access to administrative records for GQ facilities. The completeness of such records should also be assessed by GQ type.

Copyright National Academy of Sciences. All rights reserved.

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. It is important to note that "race" and "ethnicity" are social-political constructs and often associated with national origins. They are also fluid, depending in part on the format of the survey or census questions and who is the respondent (self, other household member, proxy). The measurement of race and ethnicity in the 2020 Census was complicated by several factors. While an expected change by the U.S. Office of Management and Budget of the measurement standard from the two-question format for race and ethnicity to a combined question did not occur, the Census Bureau made important changes to the existing questions. These changes, which included the provision of write-ins for the White and Black categories and expanded data capture and coding, undoubtedly played a role in the greater diversity measured in the 2020 Census.

Conclusion 10.1: The 2020 Census depicted a more diverse nation than did the 2010 Census. The White Alone population declined significantly, particularly among Hispanic people; the Hispanic population increased; and the Two or More Races population increased substantially, particularly among Hispanic people. It is unknown the extent to which these changes were due to demographic changes (e.g., more multiracial children); changes in self-identification; the provision of write-in origins (perhaps not intended to indicate racial identification) made possible by the format and processing changes to the 2020 Census race question; and other effects of the 2020 Census format and processing changes.

Conclusion 10.2: The 2020 Census had poorer quality data on race and ethnicity compared with the 2010 Census in terms of coverage error (net overcounts and undercounts), rates of missing and imputed responses, and the noise infused by the 2020 Census Disclosure Avoidance System.

- *The 2020 Census exhibited larger coverage errors compared with 2010 for race and ethnic groups. There were increases in net overcounts for some groups, particularly non-Hispanic White Alone people and Asian people, and increases in net undercounts for other groups, particularly Black people, Hispanic people, and American Indians and Alaska Natives. Consequently, differences among groups widened significantly in 2020, with adverse implications for uses of census data to allocate fixed resources, such as representation, funding, and services.*

Copyright National Academy of Sciences. All rights reserved.

- *The 2020 Census had higher imputation rates (to account for missing and inconsistent responses) for race and ethnicity compared with 2010, particularly for responses not provided by a household member via the internet, a mailed-back paper questionnaire, or telephone. Moreover, the 2020 Census had much higher imputation rates for race and ethnicity for people in group quarters compared with the already high rates in 2010.*
- *The 2020 Disclosure Avoidance System injected noise of considerable magnitude into the Redistricting File for small race and ethnicity groups in small governmental jurisdictions, such as incorporated places and American Indian tribal lands and Alaska Native villages.*

Recommendation 10.1: The U.S. Census Bureau should conduct research to determine how the changes in format and processes that were made in the 2020 Census and in the American Community Survey beginning in 2020 affected the distributions of race and ethnicity. Such research should use qualitative, quantitative, and simulation methods to ascertain: how respondents viewed and used the 2010 and 2020 formats; trends in multirace reporting by age, sex, race, and ethnicity; how samples of 2020 respondents would have been categorized using the 2010 format, data capture, and coding; and the implications of differences in write-ins by response mode (e.g., more write-ins for internet responses) on race distributions among geographic areas. The Census Bureau should communicate the results of this research to data users to assist them to understand the implications of the changes for time series, and should also use the findings to inform 2030 Census planning.

Recommendation 10.2: To improve the quality of reporting on race and ethnicity in the census, American Community Survey (ACS), and other federal data collections, the U.S. Office of Management and Budget (OMB) should revise Statistical Policy Directive No. 15 to adopt a combined, check-more-than-one race/ethnicity classification with both Hispanic and Middle Eastern or North African as main categories, in addition to White, Black, American Indian and Alaska Native, Asian, and Native Hawaiian and Other Pacific Islander. OMB should formalize any changes to the race and ethnicity reporting standards as soon as practicable, to permit the U.S. Census Bureau to implement the new race/ethnicity question in the ACS and in the 2030 Census. OMB should allow agencies to collect data only for the major categories in the combined

Copyright National Academy of Sciences. All rights reserved.

question when added detail (as in the 2020 Census) would impose undue administrative burdens. A decision by the Census Bureau on including added detail and write-in spaces, along with expanded data capture and coding procedures, for a combined race/ethnicity question for the ACS and the 2030 Census should await completion of research on such matters as how respondents viewed the 2020 question format and the effects of response-mode differences in write-in responses on diversity among geographic areas.

Recommendation 10.4: The U.S. Census Bureau should consult as early as possible with the redistricting community, which includes state legislators, redistricting commissions, political parties, and political consultants, among others, to determine the optimum set of tabulations to include on the 2030 Redistricting File, whether the current race and ethnicity categories are retained or the proposed revisions take effect. The consultation should include consideration of streamlining the number of race/ethnicity categories in the file and combining blocks with small populations (using input from localities) to maximize the ability to protect confidentiality while maintaining accuracy.

IMPACT OF DISCLOSURE AVOIDANCE CHANGES ON 2020 CENSUS DATA PRODUCTS

Regardless of unique challenges forced upon the 2020 Census, other challenges were the result of the Census Bureau's own decisions. Chiefly, 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. While confidentiality protection is a critically important responsibility of a statistical agency, this decision was made without appropriate consideration and balance regarding the utility of resulting census data products to fulfill the many important functions of census data. In short, the new DAS was not ready for use in 2020 Census production, degrading the value of 2020 Census data products in terms of both quality and timeliness. Moreover, lingering questions remain about both the simulated database reconstruction attack that motivated the new DAS and the degree of confidentiality protection that was ultimately realized through final parameter settings. This has arguably harmed the 2020 Census data products, some of which are not set to be released until late 2024, and the reputation of the Census Bureau.

Copyright National Academy of Sciences. All rights reserved.

Conclusion 11.1: In an era of Big Data, linkage technology, and other aspects of today's computational and data environment, it is difficult for the decennial census to balance the need for accurate data for small areas and small population groups with an adequate level of confidentiality protection. When evaluating methods and levels of confidentiality protection, stewardship of census data for the public good requires balancing societal needs for the information—particularly for equity in representation, fund allocation, and many other uses—with plausible harms of breaches of confidentiality.

Conclusion 11.2: The decision of the U.S. Census Bureau to respond to the threats to confidentiality protection at a very late date in 2020 Census planning with a new, more complex Disclosure Avoidance System (DAS) using differential privacy-based algorithms went counter to long-standing principles of decennial census planning. The approach had not been tested in a census environment nor had the ability of the algorithms to handle critical user data needs been assessed; the Census Bureau had no backup plan should implementation of the new DAS prove challenging. The decision to continue to deploy the new DAS in the face of serious implementation problems has resulted in marked delays in delivery of data products, with some variables and types of geographic units of questionable utility, and other variables and geographies not provided at all. In addition, it is not clear that the chosen privacy budgets for the various 2020 Census data products, with high values of the ϵ parameter that trade off accuracy with confidentiality protection, provide much actual protection.

Conclusion 11.3: Sustained, interactive communication with data users in advance of major changes to data products from a census or long-running survey is critical for a statistical agency to maintain credibility with and serve its user community. The late decision to adopt new confidentiality-protection techniques for the 2020 Census led to untimely and inadequate user communication on the part of the U.S. Census Bureau. The comment periods after the release of each demonstration file for the Redistricting and Demographic and Housing Characteristics Files were valuable, but many users lacked the staff and resources to respond adequately.

Recommendation 11.1: For 2030 Census data products, the U.S. Census Bureau should adopt the risk-utility framework recommended by the Advisory Commission on Data for Evidence Building, which accepts that disclosure risk is a continuum, that not all data items are equally sensitive, and that federal

Copyright National Academy of Sciences. All rights reserved.

statistical data need to be accessible and useful for a wide range of users and uses.

Recommendation 11.2: At a minimum, the 2030 Census should provide as-collected (i.e., unaltered) total population counts for all governmental units (states, counties and equivalents, minor civil divisions, incorporated places, recognized tribal areas, school districts) and quasi-governmental units (census county divisions, census-designated places, and tribal statistical areas), no matter how small. In addition, census block population totals should add up to block groups, census tracts, counties, and states.

Recommendation 11.3: For the 2030 Census data product plan, the U.S. Census Bureau should begin immediately on a multipronged research program with ample testing and opportunities for feedback and dialogue with the data user and stakeholder community, broadly defined. The goal should be an end-to-end plan by 2027–2028 for producing a suite of 2030 data products that serve user needs, are appropriately protected, and meet the time schedule of the 1990–2010 Censuses.

Recommendation 11.5: The U.S. Census Bureau should welcome initiatives to add language to appropriate legislative vehicles, such as the Foundations for Evidence-Based Policymaking Act of 2018, that prescribes responsibilities and penalties for data users, in addition to agency staff, for willful, harmful disclosure of confidential information.

LEARNING FROM 2020, PREPARING FOR 2030

Envisioning the 2030 Census and beyond, we echo the general guidance of previous National Academies of Sciences, Engineering, and Medicine panels. To effect major change in census conduct, the Census Bureau should 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. At the same time, there is danger in excessively 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, such as a program of earlier and sustained In-Office Address Canvassing including portions that were descoped for budgetary reasons. We encourage documenting and improving upon the contingency procedures that had to be developed and invoked extemporaneously in 2020 (e.g., enhancing

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

work with community and state/local/tribal government partners), many of which may hold promise for 2030 and beyond.

Recommendation 12.1: To prioritize its testing and development of both procedures and technical systems for the 2030 Census, the U.S. Census Bureau should take the following as primary objectives:

- **Maximize self-response to the census, including better matching of contact and communication strategies to the desired response mode, with particular attention to hard-to-reach, at-risk populations;**
- **Improve the quality of data in Nonresponse Followup, including reducing if not eliminating the use of low-quality proxy reporting when there is an alternative available;**
- **Reduce gaps in coverage and data quality associated with race, ethnicity, and socioeconomic status;**
- **Improve the quality of address listings and contact strategies for all living quarters, including group quarters; and**
- **Realign the balance between utility, timeliness, and confidentiality protection in 2030 Census data products.**

It is important to note a deliberate omission from this list of suggested priority goals. Though we encourage the Census Bureau to continue its solid work on using administrative records data as a supplement to census operations, including judicious use of records data to enumerate some nonresponding households, we do not support a movement toward an administrative records-based census in which records would be used as original enumeration for a large percentage of the population. We do not believe that the vast array of conceptual, technical, and legal issues associated with a move toward a records-based census are resolvable in time for production of the 2030 Census. It is vitally important that the goals and designs of the 2030 Census be developed in true partnership with census data users; myriad community stakeholders; and state, local, tribal, and government partners—all partnerships that make the decennial census the essential, grand civic ceremony upon which the nation relies.

Copyright National Academy of Sciences. All rights reserved.

– 1 –

Introduction

1.1 THE CENSUS AS ESSENTIAL CIVIC CEREMONY

In taking stock of the 2020 Census of the United States, it is instructive to begin by considering the stakes involved in the decennial exercise and how fundamental it is to the function of the nation. The survival of a democracy necessarily requires that successive generations of the population embrace its institutions, procedures, and practices. Since 1790, the U.S. census has been such an institution. Sometimes described as the nation's family photograph, the U.S. census is the federal government's largest undertaking during peacetime. It is a recurring civic ceremony in which everyone counts, and it reaffirms a commitment to equality among all.

The census is a cornerstone of American democracy because political representation is explicitly tied to population counts. An "actual Enumeration" is mandated in Article I, Section 2, of the U.S. Constitution, as amended by the 14th Amendment. The constitutional purpose of the census is apportioning the U.S. House of Representatives among the states, but the extent to which the decennial census reshapes the contours of political representation and power go far beyond apportionment. After each census, state and local governments create legislative districts—from U.S. Congress down to such entities as school districts—in the process of redistricting. Most states rely on decennial census data to conduct redistricting, though the use of those data is explicitly required by state constitution or statute in only a minority of states (National Conference of State Legislatures, 2021). Congress enacted Public Law 94-171 in 1975 to ensure that the U.S. Census Bureau provides relevant census data to the states for redistricting purposes, and legal requirements for equal-

Copyright National Academy of Sciences. All rights reserved.

population legislative districts currently support the expectation that census data on the voting-age population by race and ethnicity will be available for all 5.8 million inhabited census blocks in the nation.

The decennial census also serves as a vehicle for fairness in other contexts, such as the distribution of funds for federal programs. Although the annual amount varies with appropriations, in Fiscal Year (FY) 2017, 316 federal spending programs relied on 2010 Census-derived data to distribute $1.504 trillion, or about 7.8 percent of the Gross Domestic Product (Reamer, 2020). An updated analysis for FY 2021 concluded that 353 federal spending programs allocated $2.8 trillion based on decennial census and American Community Survey data (Villa Ross, 2023). These funds were distributed to various government entities, and the data inform the eligibility standards that affect the provision of benefits and the definitions of regulatory action throughout the nation. The more accurate the census count, the more equitable the distribution will be. Undercounted geographic areas or groups will lose their share of funding while overcounted areas or groups will unduly benefit.

Although the number of data items collected in the 2020 Census is limited—age, sex, race/ethnicity, relationship to householder, housing tenure—the uses of those data are myriad,[1] and the decennial census is the largest and most vital undertaking of the federal statistical system. Census data are highly valuable beyond these governmental purposes because the census is the only public geographically comprehensive database. Census data underlie government and business planning, provide the denominators for health and other vital statistics, serve as the evidence base for assessing claims of discrimination and deprivation of civil rights, and fuel countless other private and public uses. Small geographic areas, such as small towns, lack the resources to create comparable data, so they are particularly reliant on census data; the number of these small jurisdictions is considerable[2] and the census has long produced small-area data even in the precomputer age.[3] So, too, do small populations by race, ethnicity, and age

[1]See Appendix F.3 for further examples of decennial census data. The 2000 Census was the last to collect additional socioeconomic data on the population in the "long-form" sample; beginning in 2005, such items as income, education, veterans' status, and many others are collected from large population samples in the continuous American Community Survey.

[2]For example, 50% of incorporated places had no more than about 1,125 people in 2020; the corresponding figures for minor civil divisions of counties and American Indian reservations were about 825 people and 550 people, respectively (Census Geographies Project, 2022:21, 24, 27). As of 2020, the United States had 3,143 counties and equivalents, 19,483 incorporated places, 17,780 minor civil divisions, and 337 American Indian tribal governments, plus large numbers of statistical areas, including census-designated places, census county divisions, Native American areas, census tracts, block groups, and blocks (Census Geographies Project, 2022:Table 2-1A).

[3]Per Census Geographies Project (2022:Ch. 3), census publications date back to 1790 for counties and minor civil divisions, to 1880 for incorporated places (and even father back for some large cities), to 1910 for census tracts in the cities that pioneered these neighborhood areas, to 1930 for enumeration districts (the equivalent of today's block groups), and, for blocks in large cities, to 1940 (housing characteristics) and 1960 (population and housing characteristics).

Copyright National Academy of Sciences. All rights reserved.

depend on census data to monitor their sizes and distributions, as is essential for a nation striving to curb inequity.[4]

The multiple functions and high quality standards required of the census count justify the Congressional willingness to appropriate adequate funds for this vast undertaking—approximately $13.7 billion for the 2020 Census (U.S. Government Accountability Office, 2023). In return for their cooperation in this decennial civic ceremony, the people of the United States receive free public data that is presumably of high quality.

The census is an essential civic ceremony, but it has never been an easy one. Each decade brings a larger and more complex population to cover, new technologies and methods to adapt, and new difficulties with which to contend. On that score, the 2020 Census was exceptional, posing an unprecedented combination of challenges and operational decision points to census takers, not least of which was a lengthy shutdown in operations necessitated by the onset of the COVID-19 pandemic at exactly the time that 2020 Census operations were ramping up. In the truest sense of the word, the 2020 Census was extraordinary—so much so that the mere fact that data collection was completed at all remains a signature and very laudable achievement. Against this backdrop, this second and final report of the Panel to Evaluate the Quality of the 2020 Census delves into the still-open question of whether the extraordinary enumeration produced results of the quality commensurate with the task.

1.2 NATURE AND LIMITATIONS OF THIS STUDY

1.2.1 Panel Charge and Data Analysis Subgroup Structure

The Census Bureau sponsored this study with a very general but expansive statement of task:

> The National Academies of Sciences, Engineering, and Medicine will appoint an ad hoc panel to review and evaluate the quality of the data that were collected in the 2020 Census. As part of its work, the panel will:
>
> 1. Review information from the Census Bureau on the data collected as well as various process measures and indicators of data quality obtained as part of the 2020 Census operations;
> 2. Review other available information, such as results from demographic analysis, process measures and preliminary results from the post-enumeration survey; and analyses of administrative records; and
> 3. Consider the results from evaluations of similar indicators from the 2010 and 2000 Censuses.

[4]Executive Order No. 13985 (2021) created new requirements related to data equity in federal programs. The order also established an Equitable Data Working Group (2022), which issued a vision and recommendations on April 22, 2022, one of which was to "make disaggregated data the norm while protecting privacy."

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

> The panel will produce an interim report with its initial findings and conclusions, and a final report that includes conclusions about the quality of the data collected in the 2020 Census and makes recommendations for further research by the Census Bureau to evaluate the quality of the 2020 data and to begin planning the 2030 Census. The panel's reports will be reviewed according to institutional review procedures and released publicly on the National Academies Press web site.

The distinguishing characteristic of this panel, since the earliest proposal discussions with the Census Bureau, is its designation of a dedicated data analysis subgroup from within our membership. The members of this subgroup and panel staff were sworn in and conducted their work behind the Census Bureau's information technology (IT) firewall, via the Census Bureau's virtual desktop interface (VDI). In furtherance of its charge, the full panel met 10 times in plenary between July 2021 and May 2023—under the circumstances, in principally or exclusively virtual format—with the series of open meetings in 2021 dedicated to receiving overview presentations from Census Bureau staff related to the processes used for the 2020 Census. The data analysis subgroup met much more frequently—typically twice a week over most of its span—to conduct its work, and the panel's meetings in 2022 were structured to permit the subgroup to brief the rest of the membership about their work while keeping sensitive information behind the IT firewall. Full-panel members needed to be duly sworn or IT-cleared to participate in these discussions, which were conducted via the Census Bureau's VDI.[5]

We issued our interim report in April 2022 (National Academies of Sciences, Engineering, and Medicine, 2022), delayed from its original timeline because of the time needed to establish the data-access arrangements. Because those arrangements were just being secured and there was not an existing base of data for the subgroup to work with, our interim report could not reflect original data analysis. Still, the interim report laid down important markers for this report, and this work extends from the framework developed in the first report.

1.2.2 Defining Census Quality

A few broad comments are in order, to set the stage for what is and what is not included in this final report, and to calibrate expectations accordingly. The first two of these comments stem directly from the precise wording of our charge, which obliges us to "review and evaluate the quality of the data that were collected in the 2020 Census." This language requires us to be clear about what we mean by the quality of a decennial census and its data.

[5]The panel's final meeting in May 2023, to finalize concepts and language for this report, involved only data and analyses that had been cleared for release/panel use by the Census Bureau's Disclosure Review Board.

Copyright National Academy of Sciences. All rights reserved.

Box 1.1 Federal Committee on Statistical Methodology (2020) Data Quality Framework

Domain	Dimension
Utility	Relevance
	Accessibility
	Timeliness
	Punctuality
	Granularity
Objectivity	Accuracy and reliability
	Coherence
Integrity	Scientific integrity
	Credibility
	Computer and physical security
	Confidentiality

SOURCE: Federal Committee on Statistical Methodology (2020:2).

The concepts of data "quality" and "error" for a census were a major focus of our interim report, and we repeat the basic precepts from that report because they remain immediately relevant here. We held in the interim report that perfection—the absence of all error—is an unattainable and unrealistic standard for any census. We observed that the assessment of the quality of a census depends on the filters and criteria applied in the review, the relative weight put on uses of census data, and the fitness of the collected data to meet those uses.

We noted in the interim report that that there is no simple or unique scorecard for determining census quality, and this basic truth has been continually reinforced as this study has progressed. In our work, we have also continually returned to the frameworks for fundamental responsibilities and functions of statistical agencies, as embodied in such sources as our parent Committee on National Statistics' *Principles and Practices for a Federal Statistical Agency* (National Academies of Sciences, Engineering, and Medicine, 2021) and written into law as the Evidence-Based Policymaking Act of 2018 (44 U.S.C. § 3563(a)(1)). In particular, we resonate with the framework for data quality offered by the Federal Committee on Statistical Methodology (2020:2) that defines data quality as "the degree to which data capture the desired information using appropriate methodology in a manner that sustains public trust" and that decomposes data quality into 11 component dimensions, as listed in Box 1.1. It is useful to think of census quality as a composite—a calibration, if not a complete balance—of these multiple dimensions and, relatedly, to be cautious of the harmful effects of unduly weighting any one dimension above all others.

Copyright National Academy of Sciences. All rights reserved.

1.2.3 The Enduring 2020 Census

The second broad comment directly stemming from our charge to "review and evaluate the quality of the data that were collected in the 2020 Census" is that it is necessarily awkward—if not impossible—to render a complete evaluation when the circumstances and timelines of the 2020 Census are such that its detailed data products were not yet released or were still in development as this report was drafted. The major tranche of data that comprises the P.L. 94-171 redistricting files goes into deep geographic detail but utilizes only a crude age categorization (to reflect voting age) and combinations of high-level race categories. More detailed, publicly released tabulations of the 2020 Census by age, race, ethnicity, household relationship, tenure, or combinations thereof, awaited release of the Demographic and Housing Characteristics (DHC) File on May 25, 2023—as this report was in final stages of drafting and simply too late to work with other than to acknowledge the release. Of necessity, then, our work has focused on process and methods, drawing inference about the quality of 2020 Census data based on how they were collected along with other very high-level measures of quality, but with very little tangible reference to the actual census data products themselves.

A related, further complication also demands mention, which is that the delays in issuing 2020 Census data products have produced corresponding delays in the generation of the Census Bureau's internal evaluations and assessments of census operations. This virtually ensures that some discrepancies will arise between what we say here (and what we observed in the operational data) and the results that appear in the Census Bureau's finished data releases and evaluation reports; this is simply an inevitable result of analyses preceding each other based on still-evolving source files.

1.2.4 Limits on Analysis and Publication

Our interim report ended with some important commentary regarding Conclusion 4.6 in that report (National Academies of Sciences, Engineering, and Medicine, 2022:60):

> It will not be possible for this panel (or any other evaluator) to understand and characterize the quality of the 2020 Census unless the Census Bureau is forthcoming with informative data quality metrics, including new measures based on operational/process paradata, at substate levels and small-domain spatiotemporal resolution, unperturbed by noise infusion.

To properly give credit where it is due, the Census Bureau honored its commitment to provide us, through the data analysis subgroup, with access to the information that we requested. Mindful of the Census Bureau's ongoing 2020 Census processing and production, we tried to keep our data requests reasonable so as to be minimally disruptive to the Census Bureau's own work.

Copyright National Academy of Sciences. All rights reserved.

Compromise was needed on some data requests, but it is very important to note that none of our data requests were flatly denied, and the Census Bureau was very generous with its staff and time resources to fulfill our data requests and engage on technical points with our subgroup members. Relative to a scenario in which we as external evaluators would have been completely dependent on finished, fully approved Census Bureau products and reports, our situation was a good one indeed.

That said, there remain major issues and complications about what we can actually publish and report from our analyses. Any representation of nonpublic, behind-the-firewall data must be cleared by the Census Bureau's Disclosure Review Board procedures, including not just our analyses but also the work of the Census Bureau's own research staff in evaluation and assessment reports. Minimally, those procedures include stringent rounding rules for the presentation of numerical data; they may also involve the injection of statistical noise into the results, as is being done for the 2020 Census data products. We discuss the Census Bureau's approach to protecting the confidentiality of publicly released census data (what it terms its Disclosure Avoidance System [DAS]) at length in this report, particularly in Chapter 11. The Census Bureau's decision to place formally private algorithms that infuse noise in every data cell at the center of the 2020 DAS makes the DAS a hallmark issue for the quality of the 2020 Census outputs.

At the outset, it is important to note that the Census Bureau's stringency in trying to avoid *any* potential "leakage" of confidential information (however slight or indirect) necessarily creates awkwardness for an independent and external review such as ours, and it is necessarily counter to a position of fully public reproducible research. Though we have not seen a compelling argument that process indicators (e.g., census tract rates of internet or paper enumerations) or paradata (e.g., rates of different types of edits) are materially disclosive of sensitive personal information, the Census Bureau has been adamant for the 2020 Census that aggressive disclosure-avoidance techniques must be applied to such indicators. Consequently, our analyses were inherently exploratory in nature and favored detection of broad general patterns and trends; by nature, our analyses find correlation rather than causal influences. We examined subnational results when possible but bore in mind that detailed audits of operations in individual census tracts or even counties would be unlikely to be cleared by the Census Bureau for publication or would be misleading given the extent of noise injection likely required for release. We also limited our request to graphic summaries when possible, seeking clearance of just the graphic and not the data used to produce it. The analyses and results in this report are those for which we sought and obtained clearance through the Disclosure Review Board process.

We should also be clear that some limitations on what we write and analyze in this report are self-imposed. We never sought, and indeed did not want, to

Copyright National Academy of Sciences. All rights reserved.

directly query the full sets of processed returns (e.g., the Census Unedited File or the Census Edited File) from the 2020 Census, but rather relied on custom tabulations on targeted topics. The Census Bureau's editing and unduplication routines are highly sensitive and rightfully safeguarded, and we sought to understand the end effects of those routines without probing their mechanics in depth. Electronic data capture and field operational control systems in the 2020 Census undoubtedly created many avenues for inquiry—everything from audit trails of enumerator routes to resolution of "alerts" generated in the field systems to inferences from keystroke sequences and completion times in the internet questionnaire—that we did not have the wherewithal to pursue. Another reason for avoiding some of these paths is that they might provide information about the efficiency of 2020 Census processes but not necessarily the quality of 2020 Census data, which is our ultimate objective. Most notably, we did not seek nor attempt to characterize differences between publicly released 2020 Census data (i.e., the P.L. 94-171 redistricting data file) and the "true" values before application of differential privacy-based disclosure avoidance.

In the end, we have conducted as full an assessment of the 2020 Census as our limited time and resources and the parameters of the disclosure review process permitted. As presaged in the interim report, we are obliged to ask the reader to "trust us" more than would ideally be the case—to trust that we have done due diligence in the analysis, even if we cannot provide fuller results.

1.3 OVERVIEW OF THIS REPORT

We sought to make the individual chapters of this report as self-contained as possible without unduly repeating material, given content overlap between some chapters and methodological overlap in others. For convenience, our conclusions and recommendations (along with those from the interim report) are reprinted in Appendix A, and Appendix B includes an updated glossary and list of acronyms.

We begin in Chapter 2 with a brief synopsis of the design and conduct of the 2020 Census, including the array of challenges faced in the count; additional detail on 2020 Census procedures may be found in Appendix C. We then open discussion of the quality of the collected 2020 Census data with Chapter 3's examination of a phenomenon apparent from early quality measure releases: "age heaping," or an unusual extent of reported ages in round numbers ending in 0 or 5. In Chapter 4, we examine what is known from publicly released, traditional assessments of census quality: the measures of census coverage derived from comparison with the Post-Enumeration Survey and independent Demographic Analysis (additional technical detail arising from this discussion is provided in Appendix D).

Copyright National Academy of Sciences. All rights reserved.

Following this overview, in the next five chapters we explore major topics in 2020 Census operations, recapping what was planned to happen in the 2020 Census and what actually occurred due to the unique conditions in 2020, and presenting results from our analysis of operational data. In turn, these chapters cover:

- Development of the Census Bureau's Master Address File (MAF), which defines the operational frame for the count (Chapter 5);
- Provision for self-response to the census, including via the internet as the primary channel (Chapter 6);
- Conduct of the resource-intensive Nonresponse Followup (NRFU) operation to try to complete the count (Chapter 7);
- New-for-2020 use of administrative records-type data to enumerate households for which reliable external data were available and that were not successfully reached in NRFU (Chapter 8); and
- Counting of the small but very important population residing in nonhousehold or group quarters locations, such as college and university student housing and health care facilities (Chapter 9; see also the detailed code list used in 2020, reprinted for reference in Appendix E).

The Census Bureau made important changes—with vital ramifications for understanding census data quality—in two other core aspects of the 2020 Census, and these form the basis for the final two subject chapters. Chapter 10 describes the measurement of race and ethnicity (Hispanic origin) in the 2020 Census and the coding of those data, relative to past practice, while Chapter 11 focuses on the Census Bureau's extremely consequential decision to supplant its existing confidentiality-protection routines in the late stages of 2020 planning, replacing them with an untested and undeveloped Disclosure Avoidance System that affects every 2020 Census data product. (Additional material related to the disclosure avoidance topic is included in Appendix F.)

Our formal recommendations generally speak to the nascent 2030 Census planning process and are peppered throughout the main subject chapters, numbered in the sequence in which they appear in the text. With Chapter 12, we close the report by weaving all of the individual recommendations into a structured narrative—learning from the 2020 Census experience to shape the 2030 Census (and its successors), and identifying key objectives and priorities for 2030 Census testing and development.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

– 2 –

Overview of the 2020 Census

In this chapter, to set the context for the rest of the report, we describe the evolution of the design of the 2020 Census and how its procedures were planned to occur, as well as a synopsis of the operational changes made before and during the census process. We intend this chapter as an orientation to the terms and procedures described in the detailed content chapters of this report, and it is complemented by Appendix C in which we delve into some details of the operational background.

2.1 FOUR KEY INNOVATION AREAS, AND INITIAL VISION FOR THE 2020 CENSUS

The 2020 Census was designed around four "key innovation areas" that were recommended in 2011 by our predecessor Panel to Review the 2010 Census as priority topic areas in which to focus research and development. These topic areas were: (1) field reengineering, "applying modern operations engineering to census field data collection operations;" (2) self-response options, notably including provision for response via the internet; (3) administrative records, using these third-party data "to supplement and improve a wide variety of census operations;" and (4) continuously improved geographic resources, reducing the need for "once-a-decade overhauls" of the geographic databases (National Research Council, 2011:Rec. 1). The U.S. Census Bureau began work in each of these areas, in an alternating series of national content tests (mail/Self-Response only, without a field component), smaller census tests (with field Nonresponse Followup [NRFU] work), and focused testing activities (e.g., a test of partial block canvassing methods, in which listers performed address

Copyright National Academy of Sciences. All rights reserved.

list verification activities on just targeted pieces of a block rather than a complete walk of every street). Importantly, the prototype systems developed for these tests—including MOJO for operational monitoring and enumerator route optimization and COMPASS for field data collection, as well as the existing Centurion and emerging Primus platforms for internet response—were developed in-house by Census Bureau staff, under the general heading of the Reorganized Census with Integrated Technology (ROCkIT) Working Group, and their success provided important proofs of concept. In parallel, the Census Bureau also launched a Geographic Support Systems Initiative to better harness information from local-area geographic files.

The four key innovation areas formed the centerpiece of the first version of the Census Bureau's Operational Plan for the 2020 Census (U.S. Census Bureau, 2015), released in November 2015, and that plan signaled some very marked changes from traditional Census Bureau practice. Indeed, in the geographic resources area, the first operational plan clarified that the 2020 Census was already underway as of early 2015, in the form of In-Office Address Canvassing (IOAC) to identify issues with the Master Address File (MAF) without a field visit; the in-office work would permit In-Field Address Canvassing (IFAC) in 2019 to focus on a much smaller number of cases than the complete block canvassing (field staff walking every street in every block) in 2010. In the self-response area, the first operational plan signaled a commitment not only to internet response but to "Non-ID" response—permitting respondents to answer the census anywhere, from any device, without strictly requiring the respondent to enter an ID number sent by physical mail or left at the respondent's home. In terms of field reengineering and administrative records, the original operational plan noted many key decisions yet to be made but underscored the desire to reduce the number of enumerator contact attempts during the NRFU operation, including the possibility of using administrative records to enumerate households in some instances when response was not possible but sufficiently high-quality records were available. The plan also foretold a smaller field infrastructure for the 2020 Census relative to 2010, calling for 250 Area Census Offices (a halving of the number of Local Census Offices established for the 2010 Census, and consistent with the halving from 12 to 6 permanent Regional Offices that had already been implemented). A critical starting point for census operations is the delineation of type of enumeration areas (TEAs), governing the approach used for questionnaire delivery and data collection in each census Basic Collection Unit (BCU).[1] The first version of the 2020 Census operational plan called for an important simplification: the 2010 Census Update/Leave TEA (in which census enumerators would update the address list

[1]The 2020 Census adopted the new BCU nomenclature, in place of the "collection block" (as distinct from final "tabulation blocks") label used in previous censuses. Aggregated to the census tract level, the 2020 Census BCUs roughly tracked the existing census block/tract geography developed for tabulating the 2010 Census and updated over the next several years.

Copyright National Academy of Sciences. All rights reserved.

and drop off a questionnaire package at the same time, in areas without mail delivery) and Update/Enumerate TEA (in which enumerators would update the address list and conduct the household interview at the same time) were to be consolidated into a single Update Enumerate operation and TEA.

2.2 EMERGENCE OF EXTERNAL PRESSURES

Most of the major external pressures that would shape the 2020 Census took hold in the middle to late years of the decade, but one was present from the outset. The need for continuing resolutions to fund government operations became commonplace, continuing funding at the previous fiscal year's level because a new appropriations bill was not yet enacted, and doing so for ever-longer periods of time. Table 2.1 illustrates the requested and final enacted appropriations for the 2020 Census; for fiscal years 2012–2021, the length of time in which 2020 Census (and parent U.S. Department of Commerce) funding was governed by continuing resolution ranged from 48–216 days. Though Congress would exceed 2020 Census budget requests late in the decade, the opposite was true earlier in the decade. Hence, "funding uncertainty" became a familiar concern in the middle years of the decade, at the same time that the Census Bureau was working to sustain a program of 2020 Census planning and testing. This funding uncertainty would prompt the Census Bureau to scale back field testing of census operations in 2017 and 2018, and would be the cited reason for several other operational decisions, as we discuss below. Funding-related cutbacks might not be directly traceable to ultimate data quality, but they did limit the extent to which census processes and systems could be tested in advance.

The content, and not just the budgetary process of the census, came under pressure in the late stages of the decade in two important respects—both of which first came to particular prominence in December 2017 and both of which were grounded in the same provision of census law—namely the requirement that the U.S. Congress be notified of the specific questions to be asked in the census two years before the target date (13 U.S.C. § 141(f)). The first affected the race and ethnicity questions for 2020; the second reflected the presidential administration's objective to collect information on citizenship status, which drew the 2020 Census into advanced litigation.

On the first point, with the March 31, 2018, deadline looming for specifying the 2020 Census questions, the Census Bureau concluded that it needed to make a decision on the structure of the census questions on race and ethnicity (Hispanic origin) by December 31, 2017. The early- and mid-decade census content tests contemplated combining the race and ethnicity questions into a single question, as well as potentially adding a category for Middle Eastern and North African (MENA) to the classification. The Census Bureau's policy in

Copyright National Academy of Sciences. All rights reserved.

Table 2.1 Requested and Enacted Budget for the 2020 Census, by Fiscal Year, and Date of Enactment

	2020 Census (Dollars in Millions)			
Fiscal Year	Budget Request	Enacted Budget	Date of Enactment	Days Since Start of Fiscal Year
FY 2012	67	67	11/18/11	48
FY 2013	131	94	3/26/13	176
FY 2014	245	233	1/17/14	108
FY 2015	443	345	12/16/14	76
FY 2016	663	625	12/18/15	88
FY 2017	778	767	5/5/17	216
FY 2018	957	2,095	3/23/18	173
FY 2019	3,015	3,015	2/15/19	137
FY 2020	5,297	6,696	12/20/19	80
FY 2021	812	238*	12/27/20	87

* FY 2021 was fully funded through utilization of prior year balances rather than new appropriations.

SOURCE: Fontenot and Taylor (2021), adding in calculation of days elapsed between October 1 start of fiscal year and date of enactment of appropriations act.

the 2000 and 2010 Censuses was to conform to the U.S. Office of Management and Budget's (OMB) Statistical Policy Directive No. 15 on the collection of federal data on race and ethnicity, last revised in 1997. Because OMB did not act on proposals dating to the mid-2010s to revise the directive, the Census Bureau decided in December 2017 (and announced on January 26, 2018) that it would retain the two-question format for race and ethnicity in the 2020 Census and also exclude a new MENA category (Fontenot, 2018c). As discussed in detail in Chapter 10, this nonaction by OMB and subsequent design decision by the Census Bureau would prove consequential, particularly because the Census Bureau continued with other changes to the postcollection processing and coding of race/ethnicity data.

The second proposed content change for 2020 was more controversial, and ultimately was not implemented, though the controversy reverberated through a great deal of the media coverage of the upcoming census. In December 2017, a letter from U.S. Department of Justice officials requesting the addition of a question to the 2020 Census was circulated, and the request was said (in a January 10, 2018, press release[2]) to be under evaluation. On March 26, 2018, Commerce Secretary Wilbur Ross issued a decision memorandum

[2]Press Release Number CB18-RTQ.01; https://www.census.gov/newsroom/press-releases/2018/citizenship-question.html.

Copyright National Academy of Sciences. All rights reserved.

directing both that a citizenship question be added to the form and that the Census Bureau develop its capacity to derive citizenship information from administrative records sources. The decision was immediately challenged as an unprecedented and unjustified late addition to the census, and it touched off a first wave of litigation around the 2020 Census (see Box 2.1). In January 2019, the addition of the citizenship question was ruled a violation of the Administrative Procedures Act by a federal district court in the Southern District of New York, and the court enjoined the Census Bureau from asking the question. Under pressure for a decision on when to initiate printing of 2020 Census forms, the case was expedited to the U.S. Supreme Court, which upheld the injunction in June 2019 (*Department of Commerce v. New York*, 2019). In response, the administration issued Executive Order No. 13880 (2019), directing the work on deriving citizenship information from administrative records to continue. The entire episode generated fears that damage might have already been done and resulting public mistrust might suppress 2020 Census response by immigrants and vulnerable populations.

2.3 EVOLUTION OF THE 2020 CENSUS DESIGN

2.3.1 Progression of 2020 Census Operational Plans

Development and testing continued during the mid-decade, and Versions 2.0, 3.0, and 4.0 of the 2020 Census Operational Plan were issued in 2016, 2017, and 2018, respectively, on the heels of 2015's Version 1.1. Critically, each new version retained the broad-strokes design and four targeted areas of innovation and added new decisions on specific design parameters. But each new version also made some consequential alterations to parts of the design, mainly in the vein of tempering some of the major innovations.

By Version 2.0 of the operational plan (U.S. Census Bureau, 2016a), in line with the broader Census Enterprise Data Collection and Processing (CEDCaP) initiative, the Census Bureau decided to largely supplant the in-house software systems that had supported the early-decade census tests with outsourcing and recoding in a commercial programming environment. Specifically, citing perceived benefits in systems security and scalability, the Census Bureau contracted with Pegasystems Inc. in May 2016 to develop a commercial off-the-shelf software base for many of its technical systems (U.S. Census Bureau, 2016b). In the process, the Census Bureau necessarily incurred loss of time and effort in recoding extant systems.

Version 3.0 (U.S. Census Bureau, 2017) formalized a particularly acute operational casualty of the prevailing funding uncertainty surrounding the 2020 Census. As originally specified, the IOAC operation for updating the Census Bureau's MAF had two major components. Blocks would first go through Interactive Review, a quick operation in which the total number of living

Copyright National Academy of Sciences. All rights reserved.

Box 2.1 Key 2020 Census Litigation

Citizenship Question (2018–2019)

- **Issue:** Authority of the U.S. Secretary of Commerce to include a question on respondent citizenship on the 2020 Census form.
- **Outcome:** The U.S. Supreme Court ruled the Secretary had not met the standards in the Administrative Procedures Act.
- **District Court cases:** *New York v. Dept. of Commerce*, 351 F. Supp. 3d 502 (S.D.N.Y. 2019); *State of California et al., Plaintiffs, v. Wilbur Ross et al.*, 358 F. Supp. 3d 965 (N.D. Cal. 2019); *Robyn Kravitz et al., Plaintiffs, v. United States Department of Commerce*, 366 F. Supp. 3d 681 (D. Md. 2019).
- **Supreme Court case:** *Department of Commerce et al., v. New York et al.* (139 S. Ct. 2551).

Deadlines for Completing the Enumeration (2020–2021)

- **Issue:** Should the U.S. Census Bureau's April 2020 COVID-19 replan schedule be abandoned to assure release of apportionment results in December 2020, regardless of the impact of the change on the quality of the collected data?
- **Outcome:** Plaintiffs requested and received a preliminary injunction to continue the enumeration for several weeks. Supreme Court stayed the injunction on October 13, 2020, ending enumeration. Plaintiffs and defendants conducted further negotiations on apportionment releases, and reached a settlement in April 2021.
- **Citations and Statements:** *National Urban League v. Ross*, 977 F.3d 698 (9th Cir. 2020); *Ross v. National Urban League*, 592 US ____ (2020); Census Bureau statement on settlement in *National Urban League v. Raimondo* (as retitled after change in administration): https://www.census.gov/newsroom/press-releases/2021/statement-on-settlement-in-national-urban-league-v-raimondo.html.

Undocumented Immigrants and Apportionment (2020–2021)

- **Issue:** Does the U.S. President have the authority to mandate the Census Bureau to count and remove undocumented immigrants from the apportionment results for states? Does the Census Bureau have the capacity to do so?
- **Outcome:** Supreme Court dismissed the case on ripeness grounds on December 18, 2020 and in its entirety on January 15, 2021.
- **Citations:** *New York v. Trump*, 485 F. Supp. 3d 422 (S.D.N.Y. 2020); *Trump v. New York*, 592 US ____ [2020].

Census Bureau Authority to Produce the Redistricting Files under P.L. 94-171 (2021)

- **Issue:** Should the Census Bureau be required to reopen consultations with the state officials as mandated in P.L. 94-171 (and completed in 2016) before implementing a 2020 disclosure avoidance method that would not report block level population tabulations as reported in the enumeration?
- **Outcome:** District court denied writ of mandamus and preliminary injunction. P.L. 94-171 file released using differential privacy-based Disclosure Avoidance System in 2021.
- **Citation:** *Alabama v. U.S. Department of Commerce*, 546 F. Supp. 3d 1057 (M.D. Ala. 2021)

Additional Resources: https://www.brennancenter.org/issues/gerrymandering-fair-representation/fair-accurate-census/2020-census-litigation; https://www.brennancenter.org/our-work/court-cases/national-urban-league-v-ross and https://www.brennancenter.org/our-work/court-cases/new-york-v-trump; https://thearp.org/litigation/national-urban-league-v-ross/ and https://thearp.org/litigation/new-york-v-trump/; https://redistricting.lls.edu/case/natl-urban-league-v-ross/.

Copyright National Academy of Sciences. All rights reserved.

quarters in the MAF is reconciled with the count of units visible in available aerial imagery. Blocks with discrepant counts would then go into Active Block Review, in which census workers at the National Processing Center in Jeffersonville, Indiana, could use multiple imagery and data sources (e.g., local area plat maps) to actively add, delete, or alter MAF entries. Funding uncertainty prompted the Census Bureau to halt Active Block Resolution in 2017; IOAC would continue with the Interactive Review segment, at the expense of adding to the eventual IFAC workload in 2019.

With Version 3.0 of the operational plan, the Census Bureau would also take a step back from a primary distinction between Self-Response areas (with initial contact by mail) and Update Enumerate areas (with address list updating and in-person enumeration done on the same personal visit by an enumerator). In the plan, the Census Bureau reinstated an Update Leave operation and TEA, in which enumerators update the address list and leave a questionnaire package (with, for 2020, instructions to return via the internet) at households but do not conduct interviews. Most of the previously delineated Update Enumerate territory would instead be covered in Update Leave; the Census Bureau would also specify that the smaller Update Enumerate operation revert to paper questionnaires rather than electronic collection on a smartphone (U.S. Census Bureau, 2017). As discussed in Section 2.3.2, a final revision to the planned TEA structures would yield the final enumeration strategies described in Box 2.2 and Figure 2.1.

By the end of 2018, other elements of the 2020 Census design had fallen into place—among them, the February 2018 finalization of residence criteria for the 2020 Census,[3] making a change in the way "temporarily deployed" military personnel (relative to "stationed" personnel) would be counted but repeating past decades' handling of groups such as adult correctional facility inmates and college students. The final 2020 Census design (preceding the count itself) was embodied in the December 2018 release of Version 4.0 of the operational plan (U.S. Census Bureau, 2018), which made two especially profound changes to proposed operations. The first tempered the pure vision of the reengineered field operations innovation area of the 2020 Census design, restructuring the resource-intensive NRFU operation to include three distinct phases. Phase 1 would include the full optimization of enumerator workload assignments as envisioned in the previous versions of the plan, but Phase 2 would cull the temporary enumerator workforce to include only those who had been most productive in obtaining interviews in Phase 1, distributing the remaining cases among those high-yield enumerators. Phase 3, dubbed Closeout, would use the same reduced set of enumerators but focus on collecting at least some information (even if population count only) for remaining households. The

[3]The "Final 2020 Census Residence Criteria and Residence Situations" were printed in the *Federal Register* on February 8, 2018 (83 FR 5525).

Copyright National Academy of Sciences. All rights reserved.

Box 2.2 Type of Enumeration Areas and Mail Contact Strategies in the 2020 Census

Version 3.0 of the 2020 Census operational plan (U.S. Census Bureau, 2017) and the original detailed operational plan for 2020 Census geographic delineations (Decennial Census Management Division, 2018) articulated six type of enumeration area (TEA) categories for the 2020 Census. About 95% of housing units in the nation are in areas coded as TEA 1 (Self-Response [SR]), with city-style addresses (i.e., number, street, apartment number if needed), and mail delivery was used for initial contact. Most of the remainder of the nation was assigned TEA 6 (Update Leave [UL]), in which census field staff would visit households to deliver a census questionnaire packet (and update Master Address File [MAF] entries, as needed) but not conduct an interview, instead relying on the household to respond by return mail or via the internet. UL was a late addition/reinstatement to the 2020 Census plans, hence its numbering as TEA 6. The 2018 reinstatement of UL coincided with the U.S. Census Bureau's decision to conduct the 2020 Census enumeration of Puerto Rico entirely as UL, due to the devastating impact of Hurricane Maria in September–October 2017 (Fontenot, 2018a).

Functionally, TEA 4 (Remote Alaska) is a suboperation of TEA 2 (Update Enumerate) as both call for enumeration interviews and address list updates to be conducted upon initial contact. The two categories differ in their main timing, with Remote Alaska operations starting earlier than the main census (in January of the census year, in the past several decennials) to facilitate transportation access to the remote areas.

These original plans also designated a separate TEA 5 (Military) for areas containing military installations. However, Fontenot (2022c) notes that "review of administrative records provided by the military" led to the Census Bureau's determination that TEA 5 should be dissolved and incorporated into TEAs 1 and 6—that is, that housing units in military areas should be enumerated by either SR or UL.

TEA 3 (Island Areas) covers the U.S. Island Areas of American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, and the U.S. Virgin Islands. Known as the "2020 Island Areas Censuses," operations in the Island Areas were very different than the stateside census: all interviews were conducted in person and on paper, using a long-form questionnaire like that used in the stateside American Community Survey.

As we discuss in greater detail elsewhere in the report, SR in the 2020 Census was further subdivided by mail contact strategy. Based on estimates of propensity to respond online, the SR group was divided into Internet First and Internet Choice groups. A paper questionnaire was sent to nonresponding households earlier in Internet Choice areas than in Internet First areas (where the follow-up cues continued to promote online response). The contact strategies were further split by language groups, with 40 out of 50 states (plus the District of Columbia) including areas designated to receive bilingual English/Spanish invitations and questionnaires (see https://www.census.gov/newsroom/press-releases/2019/contact-strategies-viewer.html).

The address unit-level detail in the MAF Development Dataset permitted us to roughly crosswalk between 2020 Census tabulation tracts and the older-vintage tracts for which the mail contact strategies were available. We pooled TEAs 2, 4, and 6 into a single NonSR (Non-Self-Response) category and assigned each tract and county the TEA and contact-strategy label representing the majority of MAF units in the area. The small number of mismatches between the vintages of tracts are mainly associated with the late-stage dissolution of the Military TEA, parts of which were placed in UL and SR (but after the publicly issued contact strategy file). The resulting TEA/mail contact strategy map, aggregated to the county level, is depicted in Figure 2.1.

Copyright National Academy of Sciences. All rights reserved.

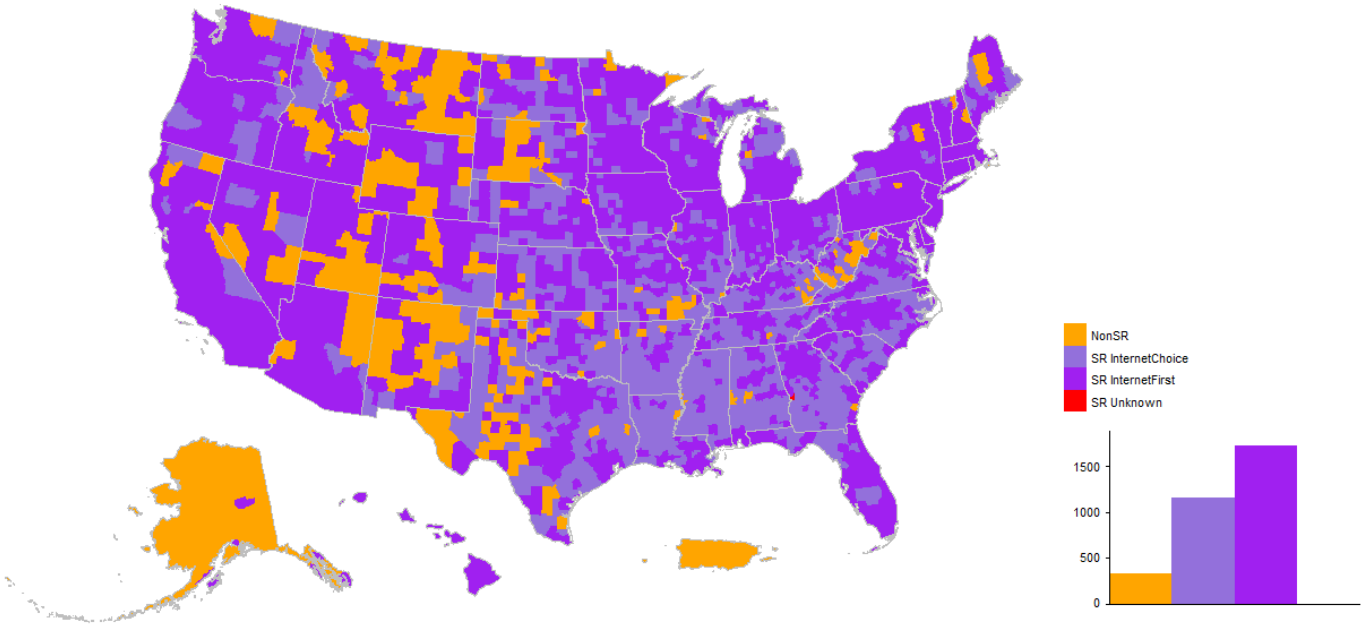


Figure 2.1 Simplified 2020 Census type of enumeration area and mailing strategy map, by county.

NOTE: SR, Self-Response, corresponds to type of enumeration area (TEA) 1. NonSR, Non-Self-Response, consolidates TEAs 2 (Update Enumerate), 4 (Remote Alaska), and 6 (Update Leave).

SOURCE: Modal TEA by tract (as embedded in Basic Collection Unit code) derived from MAF-unit level TEA designations in MAF Development Dataset; matched to publicly available dataset of mailing strategies (https://www.census.gov/content/dam/Census/newsroom/releases/2019/2020-mail-strategies-by-tract.xlsx); aggregated to county. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

second change signaled by Version 4.0 was subtly worded but was arguably the most consequential of all the precensus design changes: the new plan declared that the 2020 Census would implement a new strategy for disclosure avoidance and protecting the confidentiality of respondent information, based on the concept of differential privacy which the plan described simply as the addition of statistical noise to published tables.

2.3.2 Immediate Precensus Challenges and Operational Revisions

With Census Day 2020 fast approaching, the 2020 Census weathered additional preliminary challenges. Among these challenges, the Census Bureau raced to finalize work on the information technology support systems for the 2020 Census and to test core functionality in the 2018 End-to-End Census Test—a test that had already been heavily scaled back (from three planned sites to the single site of Providence, Rhode Island) due to funding uncertainty.

One late change, in 2019, altered a planned use of third-party, administrative records data in the 2020 Census. Prior to 2019, the Census Bureau's planned TEAs included a separate Military category, in which administrative data would play a much stronger role; "the Census Bureau had adopted the recommendation from the Defense Manpower Data Center that they could provide administrative records for the Census Bureau to use to enumerate military personnel" (Fontenot, 2021g). However, "after review of administrative records provided by the military,"[4] the Census Bureau concluded that it could not rely on the records data. Accordingly, the Military TEA covering areas with military installations was dissolved, with the previous TEA 5 being divided into TEAs 1 and 6 (Self-Response or Update Leave; Fontenot, 2022c). Similarly, military group quarters (GQs)—"barracks/dorms, treatment facilities, and military correctional barracks"—were designated for enumeration "using the methodology defined for the 2020 Group Quarters Enumeration" (Fontenot, 2021g).

In summer 2019, as census data users and stakeholders started to become aware of the sweep of the Census Bureau's new Disclosure Avoidance System, the Census Bureau committed to producing 2010 Demonstration Data Products—versions of previously published 2010 Census tabulations that subjected the 2010 data to the new differential privacy-based system slated for use in 2020. The first generation of these products was made available for rapid analysis and discussion in a December 2019 National Academies workshop (National Academies of Sciences, Engineering, and Medicine, 2020); the anomalies and unusual features of the demonstration data sparked debate over the utility of the privacy-protected data—a debate that continues into 2023 and beyond.

[4]This is described by the Census Bureau as an "enumeration test file" by Fontenot (2022c).

Copyright National Academy of Sciences. All rights reserved.

In the immediate build-up to the 2020 Census, all three of the planned major response modes to the census—internet, paper, and telephone—experienced some form of challenge:

- As described in Box 6.2, late concerns over the Pega-based internet data-collection instrument's ability to handle a massive volume of concurrent users prompted the Census Bureau to promote its in-house-developed Primus system from backup system to primary system—a decision announced on February 7, 2020. This decision was not without risk given that 2020 Census Internet Self-Response was scheduled to go live on March 12.
- In November 2017, the U.S. Government Publishing Office announced the award of the 2020 Census Printing and Mailing Contract to Cenveo, Inc.; on February 2, 2018, Cenveo filed for Chapter 11 bankruptcy protection, resulting in another several-months battle to sever the contract early. It took several more months to secure a new vendor, with R.R. Donnelley and Sons Company being awarded the new contract in late January 2019.
- Some ambitious, originally planned telephone-based operations of the 2020 Census—under the umbrella of the Census Questionnaire Assistance (CQA) operation—were considerably scaled back in 2018 and 2019. Plans for conducting some NRFU Reinterview (reinterviews for quality-assurance purposes) cases by telephone were scrapped after the 2018 End-to-End Census Test yielded fewer case completions than expected. Likewise, the plans for CQA to include a Webchat communication channel as well as a Questionnaire Status option through which CQA callers could learn "if their questionnaire had been formally received by the Census Bureau" were stricken in 2018 on cost-benefit and feasibility grounds (Fontenot, 2021j).[5]

The February 2019 omnibus appropriations act that ultimately funded most federal government operations for fiscal year 2019 included explanatory statement text that directed the Census Bureau to "devote funding to expand targeted communications activities as well as to open local questionnaire assistance centers in hard-to-count communities" for the 2020 Census.[6] In response, the Census Bureau stood up the Mobile Questionnaire Assistance (MQA) suboperation within the Internet Self-Response operation, with the intent of deploying staff equipped with mobile devices to "key locations with prominent visibility in areas with low self-response rates" (Fontenot, 2021e). In a sense, this initiative revived the partially-staffed Questionnaire Assistance Centers of the 2010 Census but without anchor to predefined physical locations;

[5]In lieu of the Questionnaire Status option, CQA phone callers were advised "to disregard mailed materials if they had already responded" to the census by other means (Fontenot, 2021j).

[6]See H.Rept. 116-9, p. 611.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

for example, a MQA site could open at a local festival or a library for a period of time, but driven by either projected low self-response rates or actual real-time response rates in an area rather than a fixed commitment. MQA staff were to be equipped with "Census Bureau-issued mobile devices" to complete interviews via the Internet Self-Response channel or to provide questionnaire or language guide assistance (Fontenot, 2021e).

2.4 THE 2020 CENSUS, PLANNED AND UNPLANNED

2.4.1 The COVID-19 Pandemic and Resulting Schedule Changes

Then, of course, the great disruptive force arrived just as 2020 Census operations were ramping up. The start of the Self-Response phase of the 2020 Census on March 12, 2020, was overshadowed by the declaration of a national emergency due to the COVID-19 pandemic on March 13. The range of state and local stay-at-home orders following the emergency declaration caused the Census Bureau to comprehensively halt all field activities on March 18—initially for two weeks, and subsequently extended—and scramble to adapt to the unprecedented change in circumstances.

On April 13, as the Census Bureau worked to gradually resume operations, Secretary of Commerce Wilbur Ross and Census Bureau Director Steven Dillingham issued a joint statement that, "in order to ensure the completeness and accuracy of the 2020 Census, the Census Bureau is seeking statutory relief from Congress of 120 additional calendar days" to deliver apportionment and redistricting data.[7] Under this proposed schedule, field offices would be "reactivated" on June 1, 2020, and census data collection would extend to October 31; apportionment totals would be delivered by April 30, 2021, and redistricting data delivered by July 31, 2021. However, in the weeks that followed, Congress did not pass any bill providing the requested statutory relief.

2.4.2 Presidential Memorandum on Undocumented Immigrant Counts, Schedule "Replan," and End of Data Collection

Before delving into the details of some of the numerous operational changes made in the 2020 Census due to the COVID-19 pandemic, we continue with a further outline of the overall chronology of events.

In this mix, a Presidential Memorandum for the Secretary of Commerce (2020) indicated the presidential administration's policy to subtract unlawful immigrants from the population base for congressional apportionment, directing that the Secretary of Commerce (via the Census Bureau) provide those state-

[7]See Census Bureau Press Release No. CB20-RTQ.16, https://www.census.gov/newsroom/press-releases/2020/statement-covid-19-2020.html.

Copyright National Academy of Sciences. All rights reserved.

level estimates at the same time as the 2020 Census apportionment counts—the statutory deadline for which had not been changed.

Under normal circumstances, the progress of a census might be marked by operational briefings and congressional oversight hearings; in 2020, it would come to be documented in court filings, discovery documents, and affidavits. Accordingly, the next steps in the chronology of events are most succinctly described in a court declaration by Census Bureau associate director Albert Fontenot, Jr.:[8]

> Once it became apparent that Congress was not likely to grant the requested statutory relief, in late July the career professional staff of the Census Bureau began to replan the Census operations to enable Census to deliver the apportionment counts by the Statutory deadline of December 31, 2020. On July 29, the Deputy Director informed us that the Secretary had directed us, in light of the absence of an extension to the statutory deadline, to present a plan at our next weekly meeting on Monday, August 3, 2020 to accelerate the remaining operations in order to meet the statutory apportionment deadline. I gathered all the senior career Census Bureau managers responsible for the 2020 Census at 8:00 a.m. on Thursday, July 30 and instructed them to begin to formalize a plan to meet the statutory deadline. … We divided into various teams to brainstorm how we might assemble the elements of this plan, and held a series of meetings from Thursday to Sunday. We developed a proposed replan that I presented to the Secretary on Monday August 3.

Secretary Ross approved and announced this "Replan" schedule that same day, under which September 30—one month earlier than the target four-month extension date previously requested by the Census Bureau—became the new target end date for data collection. This schedule prompted a new wave of litigation in several jurisdictions, arguing that the census was being unduly rushed (see Box 2.1). The case related to the speed of 2020 Census operations that progressed the furthest was *National Urban League v. Ross* (2020). Ultimately, the target end date of 2020 operations would fluctuate several times based on new plans and court injunctions, until an unsigned order of the U.S. Supreme Court related to the *National Urban League* case permitted data collection to end on October 15.

2.4.3 Changes in 2020 Census Operations Due to the COVID-19 Pandemic

The amount of disruption in census work caused by the COVID-19 pandemic and the associated field delays are difficult to overstate, and it is testament to the strength and efficiencies of the underlying census design

[8]Declaration of Albert E. Fontenot, Jr., in *National Urban League v. Ross* (2020), Document 81-1, September 4, 2020.

Copyright National Academy of Sciences. All rights reserved.

that the census could function at all. Specific changes occasioned due to the pandemic, which could ultimately affect census data quality, are peppered among publicly released operational memoranda that provided updates to the Detailed Operational Plans for individual census operations. Particularly salient examples include the following:

- Early NRFU and some administrative records enumeration for off-campus college and university students were canceled. Building off improvements realized in the 2010 Census, the Census Bureau planned to conduct Early NRFU (beginning in April rather than May) in areas in the vicinity of colleges and universities, to attempt to count students living in off-campus housing (i.e., their usual residence for census purposes). However, the pandemic-induced shutdown of operations—and the general dispersal of the college student population—led to Early NRFU being descoped (Fontenot, 2021d). In late summer of 2020, the Census Bureau mounted a contact effort to colleges and universities specifically to try to account for students living off campus who might not have been otherwise counted. "This unplanned operation resulted in the creation of additional AdRec [administrative records-based] enumeration data that was generated outside of the planned AdRec process and incorporated into the final processing" (Fontenot, 2022b).
- The original plans for enumerating specific places and facilities, which had been brokered in the Advance Visit operations for GQs and transitory locations, were scuttled by the field shutdown and scheduling delays. In the case of transitory locations, a second round of Advance Visit was scheduled but with some differences; notably, local Area Census Offices were made responsible for determining whether carnivals, circuses, or fairs would be operational during Enumeration at Transitory Locations fieldwork in September 2020, rather than contacts being brokered by calls from the Census Bureau National Processing Center's call center as done in the initial Advance Visit (Fontenot, 2021b).
- The training for NRFU enumerators was restructured and rescaled to minimize in-person contact. Plans had already called for use of online training modules between the initial orientation day (meant to be in-person/classroom) and a capstone training day (meant to include in-person trial interviewing), and this continued. But both endpoints in the training program were restructured. "The orientation classroom training day was reduced from a full day to a 2-hour appointment. The remainder of the orientation day and the capstone training day were conducted through self-study reading, a podcast, and conference calls. . . . Interactive scenarios for learning how to handle difficult and unusual situations that are a major component of the capstone did not happen in-class, [but instead] as part of the telephone capstone day" (Fontenot, 2021d).

Copyright National Academy of Sciences. All rights reserved.

- COVID-19 pandemic operations had major implications for the Update Leave and Update Enumerate types of enumeration prevalent in rural areas without mail delivery, over and above the basic difficulties of access (e.g., to American Indian lands) created by pandemic-related health orders. For instance, the address list "update" part of Update Leave was converted to a no-contact, limited practice. As planned, Update Leave enumerators were expected to make a contact attempt at housing units to ascertain address/housing unit information, in addition to dropping off a questionnaire package. As a result of the pandemic, the contact attempt was eliminated and "any updates to an address were conducted by observation only," so enumerators could identify hidden or unseen housing units in plain sight (Fontenot, 2021f). This is consequential because both Update Leave and Update Enumerate areas were excluded from IFAC eligibility, making the contact attempt a key MAF-record update activity. In the even more rural Update Enumerate and Remote Alaska operations, 2020 plans eschewed a reinterview-based quality-assurance operation in favor what was perceived as a simpler fix—deploying enumerators to work in pairs, partially as a check on each other. However, due to the pandemic, it became optional rather than mandatory for Update Enumerate staff to work in pairs—while also requiring at-least-six-feet social distancing during interviews (Fontenot, 2021i).[9]
- Pandemic-era procedures called for added functionality in the 2020 Census' telephone-based operations. As an "unplanned contingency," the Census Bureau permitted that "enumerators in areas with high COVID rates, impacted by natural disasters, or with low completion rates were allowed to use telephone follow-up" to make some of their contacts with NRFU cases by telephone, using phone numbers extracted from the Census Bureau's Contact Frame. These phone contacts would count towards determining eligibility for proxy or administrative records enumeration availability (Fontenot, 2021d). Similarly, the functionality of the CQA operation was expanded to give callers to CQA phone lines (in all languages) the option to request a callback if staff could not immediately answer; CQA representatives could return calls and either provide assistance or complete an interview, "using government furnished

[9]In changes described as unrelated to the COVID-19 pandemic, the Enumeration at Transitory Locations (ETL) operation was twice reworked to replace a field-based ETL Reinterview for quality control purposes. Original plans called for ETL Reinterview to be done by phone, but projected response rates were deemed too low. Instead, 2020 ETL interviewing was done (using paper questionnaires) by small groups of enumerators, just as was the original plan for Update Enumerate and Remote Alaska—the theory being that "the use of multiple staff for enumeration largely prevents the opportunity for the data falsification that reinterview is intended to identify" (Fontenot, 2021b).

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

equipment at home" if necessary. CQA support was also added to field calls from administrators of GQs to help them complete questionnaires, given the delay-induced need to broker new agreements for acquiring GQ information (Fontenot, 2021j). Curiously, though, phone contact was eliminated as an option in the Office-Based Address Verification and manual/clerical processing of Non-ID census returns, as part of reconfiguring operations and ensuring that resources were available for manual matching and geocoding when needed (Fontenot, 2021c).

- A subtle innovation was initially planned for the 2020 Census, in which NRFU cases in large, multiunit structures would be grouped in assignments. In these "Manager Visit" cases, enumerators would contact property and building managers first, to assess general housing unit status (and hence potentially reduce workload). However, while trying to stand up NRFU operations amidst the work-delay difficulties, it was found that "multiunit cases were not grouped correctly" in the technical systems and that, in particular, NRFU "cases that had administrative records [available] were placed in a separate workload" and thus excluded from the multiunit groupings. Hence, Manager Visits were descoped and the individual housing units within the structures released into the general NRFU workload (Fontenot, 2021d).
- The MQA program established in response to Congressional interest was premised on deployment at large public events—which were now ruled out by pandemic-era restrictions on gathering. Hence, MQA was revised in scope to focus on "sites in places where people visit when leaving home" as opposed to larger public gatherings (Fontenot, 2021e).
- Under the broader Count Review operation, which garners inputs from state demographer participants in the Federal-State Cooperative for Population Estimates (FSCPE) program, the 2020 Census was planned to involve a Post-Enumeration Group Quarters Review session in the very late stages of Group Quarters Enumeration in the field. This session, also dubbed Review Event 2, was to include both remote-access and on-site components, permitting FSCPE participants to review initial results for GQs in their states and suggest last-minute additions. Originally canceled and ultimately rescheduled as a remote-only event on September 23–28, 2020, Review Event 2 was also greatly reduced in scale; "the opportunity to add new GQs" and access to control systems and geographic databases were not provided "because once the initial review was canceled, access to the census network was removed and couldn't be reinstated in time for the modified review." Moreover, review access (in Excel spreadsheet extracts) was limited solely to the subset of "noneumerated GQs that had an enumeration disposition reason of 'Cannot Locate in Block'" (Fontenot, 2021a). Had the GQ review event been conducted as planned, it is quite possible that some major problems that affected measurement

Copyright National Academy of Sciences. All rights reserved.

of the GQ population might have been detected in time for corrective action.

- Internal Census Bureau subject-matter expert review of incoming 2020 Census data was only slated to begin with Decennial Response File 2, after some amount of processing; however, the operational fieldwork delays and resulting compressed timeline prompted the Census Bureau to start the internal review with the basic-compilation Decennial Response File 1 (Fontenot, 2021a). The COVID-19 pandemic was also said to be the cause of the cancellation of one planned study—Evaluating the Effect of the Decennial Census on Self Response to the American Community Survey (ACS)—in the 2020 Census program of evaluations and experiments (Fontenot, 2021k).
- On August 12, 2020, use of Best Time to Contact probabilities—estimated from administrative records and input into the optimizer for determining NRFU enumerator assignments—was discontinued because "using that parameter slowed the assignment of cases to field staff" (Fontenot, 2021d), and time was at a premium due to the compressed schedule.

2.4.4 2020 Census Data Products

After resolving the question of when 2020 Census field and Self-Response operations would end, the Census Bureau turned its attention to the processing of results and to conducting the even-more-delayed Post-Enumeration Survey—both of which proved more problematic and time consuming than expected. In particular, the processing encountered numerous anomalies—not unlike snags encountered by previous censuses but difficult to handle in a high-pressure environment. Meanwhile, during fall 2020, pressure to produce the estimated counts of undocumented immigrants called for by the Presidential Memorandum of July 2020 continued to bubble, and would do so until a January 2021 U.S. Department of Commerce Office of Inspector General memorandum shone light on those pressures.[10] On December 30, the Census Bureau formally acknowledged that it would be unable to satisfactorily produce apportionment totals by the year-end deadline but that it would work to do so as soon as possible. The Presidential Memorandum of July 2020 and the predecessor Executive Order No. 13880 were both formally rescinded by executive order on January 20, 2021, and the Census Bureau ultimately released 2020 Census apportionment counts and redistricting data on April 26 and August 12, 2021, respectively.

[10]See the January 12, 2021, memorandum at https://www.oig.doc.gov/OIGPublications/OIG-21-019-M.pdf. The memorandum acknowledged whistleblower reports from within the Census Bureau stating that the undocumented immigrant estimates had been cast as "a number one priority," to be complete on January 15, 2021, ahead of the change in presidential administrations.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

2.5 CONCLUSION

This brief synopsis of just some of the challenges involved in conducting the 2020 Census is important for understanding the analyses of specific operations and concepts that we pursue in the rest of this report. But it is a powerful story in its own right, one that deserves to be told in even fuller detail. While it is true that the details can be gleaned from examination of dozens of individual memoranda, we think it important to the Census Bureau's own institutional memory that these details be documented and preserved, to facilitate the derivation of lessons learned for the 2030 Census and beyond. Accordingly, we recommend:

> **Recommendation 2.1: In addition to completing its own program of evaluation and assessment reports, the U.S. Census Bureau should complete and publish a comprehensive procedural history of the 2020 Census.**

Copyright National Academy of Sciences. All rights reserved.

– 3 –

Age Heaping as an Indicator of Data Quality

This chapter presents our assessment of a classic measure of quality for the data collected in the 2020 Census—namely, the accuracy of data on age. Such accuracy may be analyzed by examining the census counts only or by comparing census age data with age data from Demographic Analysis (DA), which is based largely on vital statistics. We pay special attention to how the accuracy of age data varies by response mode. This chapter serves as a foundation for subsequent chapters that dig deeper into details for specific census operations and how they affected data accuracy, in terms of such measures as population coverage and missing responses. Chapter 11 examines the quality of the released data products and specifically the impact of the new Disclosure Avoidance System (DAS) on accuracy, timeliness, coherence, and comparability of those products.

3.1 AGE HEAPING AS A QUALITY INDICATOR FOR THE 2020 CENSUS

Examining age data is a classic method for evaluating the quality of a census (Hobbs, 2004; Jensen et al., 2023). In common with most censuses today, the 2020 Census asked for both household members' ages in completed years and for date of birth. This double approach to age measurement is helpful in editing—for example, for people with birthdays near the census date or who complete the Post-Enumeration Survey (PES).

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Age heaping refers to the population reporting certain ages at the expense of other ages, and it is common in both censuses and surveys (Pullum, 2006). In the United States, the preference is for ages ending in 0 or 5. (In other countries, different digits may be preferred; see Poston and Bouvier, 2017:290). Higher levels of age heaping indicate inaccuracy in reporting, and, for this reason, age heaping is a commonly used measure of data quality (Pardeshi, 2010; Pullum, 2006). Age heaping has been well documented, especially in developing countries, in which some respondents may be unsure of their own exact ages (see, e.g., Lyons-Amos and Stones, 2017).[1] Evidence of age heaping in the U.S. census goes back more than 150 years, although not to the degree seen in developing countries and generally diminishing over time (Author unknown, 2015; Zelnik, 1961).

Age heaping is important for two reasons. First, some uses of census data call for detailed age distributions at the state and local level that are not available from DA (which, as shown below, exhibits minimal age heaping). Also, age heaping may be an indicator of other data-quality problems in the census, affecting both coverage and other characteristics.

A March 2022 post on the U.S. Census Bureau web site reported "significant age heaping . . . in the 2020 Census results" (Jensen and Kennel, 2022). The post presented a graph of net census coverage by single years of age—essentially the 2020 line in Figure 3.1. That line shows clear evidence of heaping at multiples of 5, ranging from 20–80 years of age. The net count at each of those ages is higher than the counts for the four surrounding ages. Moreover, the extent of age heaping appears markedly more pronounced in the 2020 Census than in 2010. Similarly, West et al. (2005) document age heaping in the 2000 Census that is comparable in magnitude to 2010 and much less pronounced than in 2020.

Although graphs illustrating age heaping, such as Figure 3.1, are visually compelling, it is hard to characterize the amount of heaping from such graphs, especially at ages for which net coverage is rising or falling quickly, for example, for children under the age of 10. Also, analysis of counts for the full census is not informative about sources of age heaping. Using additional information provided by the Census Bureau, this chapter analyzes age counts in 2010 and 2020 by the type of age data reported (e.g., age only or both age and date of birth) and the mode of data collection (e.g., internet, proxy respondent).

We conclude the chapter with comments on the evidence of widespread data-quality problems (including but not limited to age heaping) in Nonresponse Followup (NRFU), particularly in proxy responses. Given that self-response rates and therefore NRFU rates are not uniformly distributed geographically or demographically, this finding has implications for equitable coverage of population groups and geographic areas.

[1]For a recent analysis in India, see https://pgbhat.wordpress.com/2020/10/31/1767/.

Copyright National Academy of Sciences. All rights reserved.

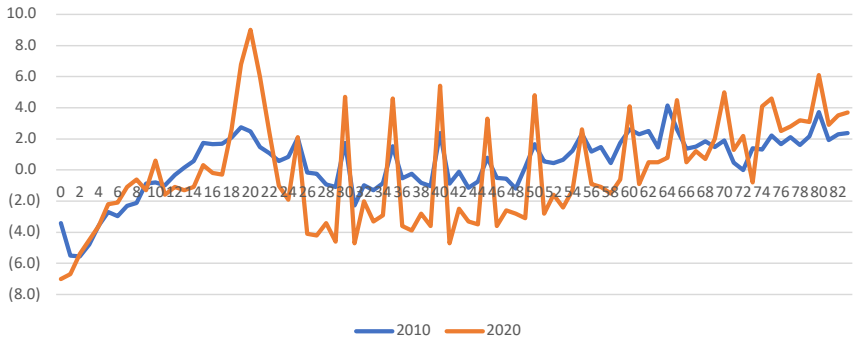


Figure 3.1 Demographic Analysis estimates of percentage net under/overcount, by single years of age, 2010, 2020, middle series.

SOURCE: 2020, middle series from Table 1—Demographic Analysis Net Coverage Error Estimates by Single Year of Age, Sex, and Series: April 1, 2020 at https://www.census.gov/data/tables/2020/demo/popest/2020-demographic-analysis-tables.html. 2010, calculated from revised middle series in Total Resident Population by Age and Race table at https://www.census.gov/data/tables/2012/demo/popest/revised-2010-demographic-analysis-estimates.html.

3.2 ANALYSIS OF AGE REPORTING IN THE CENSUS

To conduct a more thorough analysis of age heaping, we requested (and the Census Bureau provided) national-level counts by single years of age from 0–84 for the full 2010 and 2020 Censuses and for various subsets defined by the way age was reported or derived and by response mode. The Census Bureau applied privacy protection to these data by rounding (for 2010) or addition of discrete Gaussian noise (for 2020). The added noise was small compared to the magnitude of the census numbers at the national level. The source file for the data was the Census Edited File, which includes imputations for nonresponse.

Our analysis uses an age heaping metric that focuses on the counts near multiples of 5 years. It also adjusts for known patterns in the true age distribution on Census Day using the Census Bureau's DA estimates, by single years of age, for the total population for 2010 and 2020 (U.S. Census Bureau, 2012c, 2022d). The panel's metric, $\hat{h}_a$, for values of a that are multiples of 5 can be interpreted as the percentage of true ages $a - 1$, $a - 2$, $a + 1$, and $a + 2$ that get rounded to a when reported (see Box 3.1 for technical details). While $\hat{h}_a$ loses this interpretation for values of a that are not multiples of 5, one can still compute $\hat{h}_a$ for all values of a to create the line graphs that follow. This metric can be erratic in an age range, for reasons unrelated to age heaping, if the profile

Copyright National Academy of Sciences. All rights reserved.

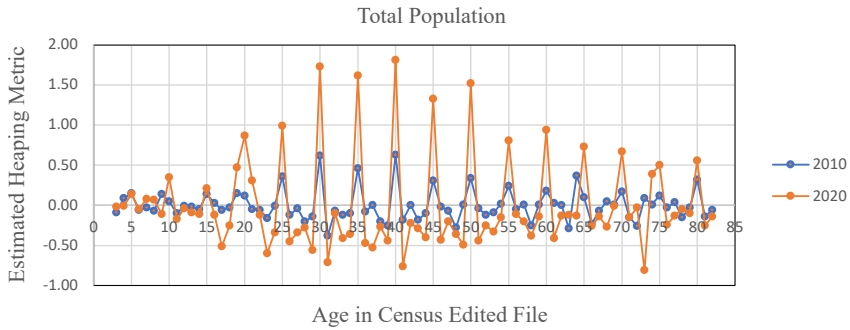


Figure 3.2 Age heaping in the 2010 and 2020 Censuses, using the $\hat{h}_a$ metric.

SOURCE: Calculations by the panel from Census Bureau tabulations from the 2020 Census Edited File. See Disclosure Review Statement; CBDRB-FY23-0180.

of true ages for the group under study differs too much from the pattern of DA estimates for that range (see Box 3.1, Step 6). This behavior can be observed around ages 17–18 when enumeration methods abruptly shift for many young adults who attend college and around the age corresponding to the beginning of the baby boom era (73–74 in 2020).

Figure 3.2 graphs $\hat{h}_a$ for $a = 3, \ldots, 82$ for the full census population in 2010 and 2020. For 2020, the most prominent feature is consistent peaks at every multiple of 5 years, with the highest peaks in the age range 30–50, declining gradually for ages above 60. There is much weaker age heaping for children. Apart from the clear evidence of heaping at multiples of 5, the most prominent feature for 2020 is a sharp rise between ages 73 and 74. The other main phenomenon is a wide peak for the range 19–21 years that seems to reflect an overcount of the college-age population on top of any heaping at age 20.

The large overcounts shown for the college-age population likely resulted from duplicate enumerations due to colleges sending students home at the start of the COVID-19 emergency declaration, just at the time when follow-up with college students was scheduled to begin. Many students were likely enumerated both at their college residence, where the Census Bureau endeavored to count them according to census residence rules at a later stage of the enumeration (see Chapter 9), and as household members reported by their parents.

The line for 2010 is similar to that for 2020 with two noteworthy distinctions. First, an atypical rise from ages 63–64 in 2010 reemerges as a more dramatic rise from ages 73–74 in 2020 (see discussion below). Of more

Copyright National Academy of Sciences. All rights reserved.

Box 3.1 Derivation of an Age Heaping Metric

1. To assess age heaping for a specific target subset of the census (perhaps the full census), for ages $a = 0,1,\ldots,84$, let r_a equal the count of persons in the target subset with *reported* age a (after any heaping), and t_a equal the count of persons in the target subset with *true* age a (before any heaping). The $\{t_a\}$ values are unobserved, but they are central to development of a heaping metric. (See paragraph 6 for a method to impute the $\{t_a\}$ values.) If r_a is substantially greater than t_a for a, a multiple of 5, that provides evidence of heaping at age a.
2. It is useful to define the analogous terms

$$\begin{aligned} R_a &= r_{a-2} + r_{a-1} + r_a + r_{a+1} + r_{a+2} \text{ and} \\ T_a &= t_{a-2} + t_{a-1} + t_a + t_{a+1} + t_{a+2}. \end{aligned}$$

3. To develop heaping metrics, one can use a simple model with the following assumptions: (1) for each age a that is a multiple of 5, true ages of $a-2$, $a-1$, $a+1$, and $a+2$ are rounded to a with probability h_a; and (2) all other true values are reported correctly. These two assumptions imply that for any age a that is a multiple of 5, $R_a = T_a$.
4. This model allows writing reported counts $\{r_a\}$ as functions of true counts $\{t_a\}$ and heaping parameters $\{h_a\}$. Specifically, when a is a multiple of 5,

$$E(r_a) = t_a + h_a(T_a - t_a).$$

5. Replacing T_a with R_a (see paragraph 3) and $E(r_a)$ with r_a allows estimating h_a by

$$\widehat{h}_a = 100\frac{r_a - t_a}{R_a - t_a}, \tag{3.1}$$

where multiplying by 100 converts the heaping metric from a proportion to a percentage. The key step is specifying a value for the unobservable t_a to plug into Equation 3.1.
6. Demographic Analysis (DA) estimates of the national age distribution can be used to impute a value for t_a. Let

$$\begin{aligned} d_a &= \text{DA estimate of U.S. persons with reported age } a\text{, and} \\ D_a &= d_{a-2} + d_{a-1} + d_a + d_{a+1} + d_{a+2}. \end{aligned}$$

Specifically, assume that true counts for the target subset are proportional to DA estimates in the neighborhood of a and substitute $t_a = d_a(T_a/D_a)$ in Equation 3.1. The assumption will be satisfied if the ratios $(t_{a-2}/d_{a-2}),\ldots,(t_{a+2}/d_{a+2})$ are all equal for ages between $a-2$ and $a+2$ for the target subset but also, more generally, if they vary linearly for ages between $a-2$ and $a+2$. Local nonlinearities in the neighborhood of a may bias $\widehat{h}_a$.

Copyright National Academy of Sciences. All rights reserved.

importance, heights of the peaks are consistently higher in 2020 than in 2010, typically by a factor of at least 3.0.

In its *Demographic Yearbook 1955*, the United Nations chose Whipple's index to evaluate the quality of age data (United Nations, 2007:5), and it is widely used (Spoorenberg, 2009). Whipple's index of age heaping is the simplest measure of age data quality because it is based only on the enumerated age distribution. The original Whipple's index is obtained by summing the number of persons in the age range between 23–62, and calculating the ratio of reported ages ending in 0 or 5 within that age range to one-fifth of the total sample.[2] This result is then multiplied by 100 to convert to a percentage. Without age heaping, this index takes the value 100. Deviation from this number indicates some degree of inaccuracy.

We use a modified Whipple's index (MWI), which subtracts 100 from the original index and divides the result by 4. Consequently, MWI = 0 indicates no heaping and MWI = x corresponds to x percent of ages being rounded to the nearest multiple of 5.[3] This property gives MWI the same interpretation as the heaping metric $\hat{h}_a$ but applied to the age range 23–62.

Table 3.1 shows MWI values for 2010 and 2020, broken down by sex and racial-ethnic group. Computed MWI values for 2010 and 2020 are 0.56 and 1.48, respectively.[4] Their ratio suggests that there was roughly 2.6 times as much age heaping in 2020 as in 2010. Using the same data, Velkoff and Jensen (2022:7) concluded "age heaping in the 2020 Census is more pronounced than in 2010."

However, MWI does not account for the possibility that the proportion of true ages in the population that are multiples of five differs from one-fifth. Indeed, when MWI is computed for the DA estimates of the age distribution, it yields values of 0.17 for 2010 and 0.14 for 2020. Consequently, substantial parts of the 0.56 and 1.48 might be due to "heaping" in the population's true ages rather than heaping by respondents. If so, that would suggest an even higher ratio going from 2010 to 2020.

Table 3.1 shows that the amount of age heaping varied by race-ethnicity and sex in both the 2010 and 2020 Censuses. Estimates of age heaping were roughly 50% higher for Hispanic and Black people than for White non-Hispanic people in both years. Notably, estimated age heaping for Asian people was the highest

[2]Using Whipple's age range of 23–62 rather than expanding it has advantages for analysis of age heaping in the 2010 and 2020 Censuses. Starting at age 23 works well because it avoids problems for the index from the apparent overcount of college students in the 2020 Census. For the upper end of the age range, Whipple likely stopped at age 62 because of lower life expectancy 100 years ago in developing countries. For recent U.S. censuses, one could easily raise the upper limit to age 72 or 82. Stopping with age 62, however, avoids the problems for the index of including pre-baby boomers in the 2010 and 2020 analyses.

[3]To convert back to the original Whipple's index, use the formula 100 + 4(MWI).

[4]A comparable value for ages 23–62 in the 2000 Census is 0.50 (calculated from West et al., 2005:3661). The unmodified Whipple's index in West et al. (2005) is 102, which converts to 0.50 after subtracting 100 and dividing by 4.

Copyright National Academy of Sciences. All rights reserved.

Table 3.1 Overall Age Heaping (Ages 23–62): Modified Whipple's Index (MWI) by Sex and Racial-Ethnic Group, 2010 and 2020 Censuses

	2010			2020		
	Male	Female	Total	Male	Female	Total
Census Total	0.63	0.49	0.56	1.64	1.32	1.48
Racial-Ethnic Group						
White Alone, Non-Hispanic	0.54	0.43	0.48	1.38	1.15	1.27
Hispanic or Latino	0.84	0.58	0.71	2.31	1.70	2.00
Black Alone	0.80	0.64	0.72	1.99	1.66	1.81
Asian Alone	0.82	0.71	0.76	1.62	1.29	1.44
AIAN Alone	0.60	0.49	0.54	1.50	1.39	1.44

NOTE: AIAN, American Indian or Alaska Native.

SOURCE: Calculations by the panel from tables available at data.census.gov for 2010 and 2020: Tables PCT12 (Census Total), PCT12I (White Alone, Non-Hispanic), PCT12H (Hispanic), PCT12B (Black Alone), PCT12D (Asian Alone), and PCT13C (AIAN Alone).

of any group in 2010, but a bit below average in 2020. MWI values were higher for males than for females, typically by 20–30%, for each combination of year and race.

3.3 SOURCES OF AGE HEAPING—TYPE OF RESPONSE

The 2010 and 2020 Census forms asked for each resident's age and date of birth (DOB). Not everyone answered these questions fully or consistently. The possible response categories are listed in Box 3.2.

Table 3.2 presents MWI values for 2010 and 2020, broken out by whether age or DOB or both were reported and, if not, by how age was imputed. The story is very simple for 2020, when age heaping was essentially a result of age-only responses. The MWI for age-only responses, 25.75, is 100 times the MWI of 0.23 for all other responses. Despite constituting less than 5% of the full census, data not shown indicate that age-only responses contributed an estimated 85% of the heaped ages in the 23–62 year age range in 2020. Results are similar for 2010 except that the MWI for age-only responses is 12.20, a bit less than half the value in 2020, and age-only responses contributed only an estimated 57% of the heaped ages in the 23–62 year range.

Figures 3.3 and 3.4 show age heaping for people providing age only and everyone else. (Note the vastly different scales in Figures 3.3 (−15 to +35) and 3.4 (−0.80 to +0.80).) The age profiles for the age-only responses (Figure 3.3) are remarkably similar for the two censuses—modest heaping through age 15, a quick increase in heaping from ages 20–30, maximum heaping through age 50,

Copyright National Academy of Sciences. All rights reserved.

Box 3.2 Response Categories for Age and Date of Birth, 2010 and 2020 Censuses

- Reported age only—random date of birth (DOB) imputed
- Reported DOB only—age calculated from DOB
- Reported both age and DOB with consistent responses—no action needed
- Reported both age and DOB but responses were inconsistent—response chosen consistent with the person's relationship to the householder
- Did not report either item—in 2020 only, administrative records (e.g., the 2010 Census, the Social Security Administration Numident File), when available, were used to fill in missing age items ("assigned") or to enumerate the whole household ("enumerated")
- Did not report either item and records unavailable—age values imputed using hot-deck method of taking reported values from a near neighbor (item imputation)
- No one in the household reported either age item or any other item and records not available—age and other values imputed using hot deck method for the entire household (whole-household imputation)

Table 3.2 Overall Age Heaping (Ages 23–62): Modified Whipple's Index (MWI), 2010 and 2020 Censuses, by Age/Date of Birth (DOB) Reporting Status

	2010			2020		
	Population			Population		
Census/Response Category	Millions	Percent	MWI	Millions	Percent	MWI
Census Total	308.7	100.0	0.56	331.4	100.0	1.48
Age Only	7.7	2.5	12.20	14.8	4.5	25.75
All Except Age Only	301.0	97.5	0.25	316.7	95.5	0.23
Age and DOB, Consistent	275.5	89.2	0.24	272.8	82.3	0.23
Age and DOB, Inconsistent	4.5	1.5	−0.30	0.9	0.3	1.16
DOB Only	4.2	1.4	−0.66	1.5	0.5	−0.73
Administrative Records Assigned/Enumerated	N.A.	N.A.	N.A.	10.7	3.2	0.19
Item Imputation (Hot Deck)	11.0	3.6	0.36	21.9	6.6	0.17
Whole-Household Imputation	5.8	1.9	1.27	8.8	2.7	0.33

NOTES: See Box 3.2 for definition of response categories. MWI, modified Whipple index; N.A., not applicable. Population values include all ages; MWI values use data for reported ages 23–62 only; negative MWI index values are not readily interpretable but only appear for small subsets of the population (those reporting DOB only in 2020 and 2010, and those reporting an inconsistent DOB and age in 2010).

SOURCE: Calculations by the panel from Census Bureau tabulations from the 2020 Census Edited File. See Disclosure Review Statement; CBDRB-FY23-0180.

Copyright National Academy of Sciences. All rights reserved.

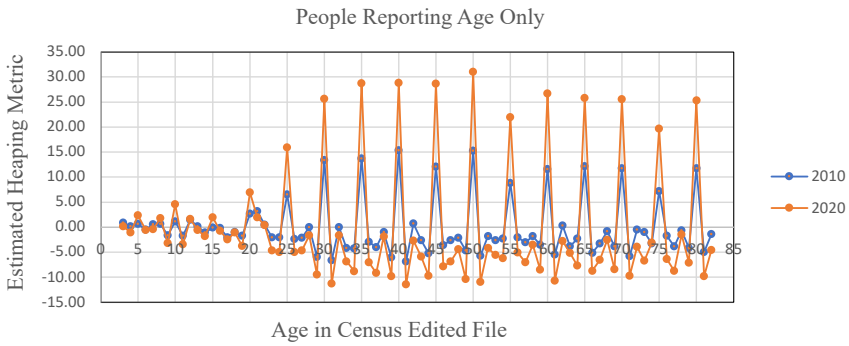


Figure 3.3 Age heaping for respondents with reported age only, 2010 and 2020 Censuses, using the $\hat{h}_a$ metric.

SOURCE: Calculations by the panel from Census Bureau tabulations from the 2020 Census Edited File. See Disclosure Review Statement; CBDRB-FY23-0180.

and slightly less heaping for ages 55 and above. At each age, the heaping metric for 2020 is roughly twice that for 2010.

In contrast, there is very little heaping displayed in Figure 3.4 for the cases that exclude age-only reporting. Except for the peak at age 20 in 2020, none of the values at multiples of 5 years exceed 0.25. Focusing on this sample, in which heaping is minimal, highlights positive outliers at age 64 for the 2010 Census and age 74 for the 2020 Census. This 10-year pattern of related outliers corresponds to people born between April 2, 1945, and April 1, 1946—roughly the last 12 months before the start of the baby boom era—though it is unclear why the DA estimates and census results fail to match better for this age.

Figure 3.5 is an age heaping graph for whole-household imputations in 2010 and 2020. The blue line illustrates consistent heaping in 2010 for all multiples of 5, from age 25 through age 70. In contrast, there is scant evidence of heaping in 2020, perhaps because of using a donor pool for the imputations that was chosen to mirror the age distribution without heaping, although the actual reason is unknown.

3.4 AGE HEAPING BY RESPONSE MODE

Table 3.3 presents MWI values for 2010 and 2020, broken out by response mode. Self-responses in both censuses exhibited modest amounts of age heaping. For example, Figure 3.6 shows results for the primary Self-Response options, paper in 2010 and internet in 2020, using the $\hat{h}_a$ metric. While there are

Copyright National Academy of Sciences. All rights reserved.

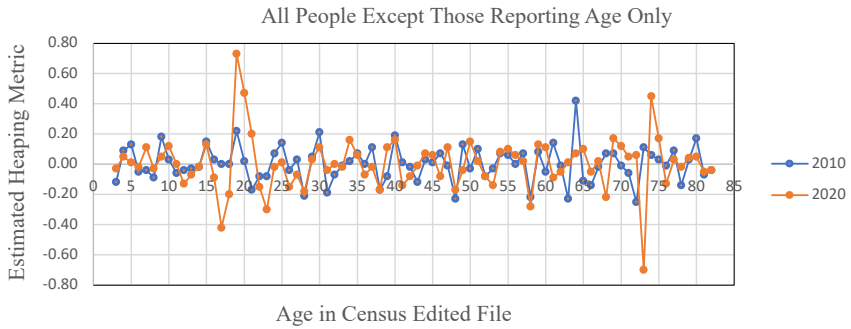


Figure 3.4 Age heaping, all respondents except those with reported age only, 2010 and 2020 Censuses, using the $\hat{h}_a$ metric.

SOURCE: Calculations by the panel from Census Bureau tabulations from the 2020 Census Edited File. See Disclosure Review Statement; CBDRB-FY23-0180.

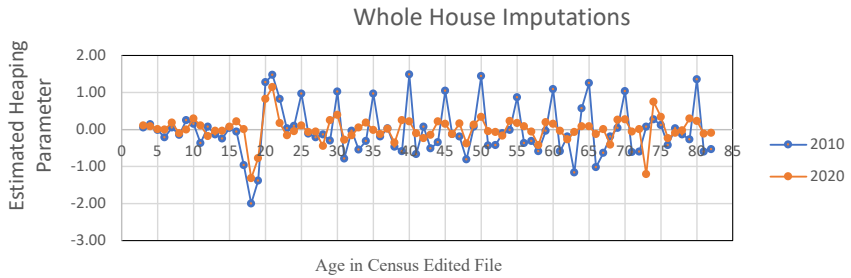


Figure 3.5 Age heaping, responses based on whole-household imputations in 2010 and 2020 Censuses, using the $\hat{h}_a$ metric.

SOURCE: Calculations by the panel from Census Bureau tabulations from the 2020 Census Edited File. See Disclosure Review Statement; CBDRB-FY23-0180.

Copyright National Academy of Sciences. All rights reserved.

Table 3.3 Overall Age Heaping (Ages 23–62): Modified Whipple's Index, 2010 and 2020 Censuses, by Response Mode

	2010			2020		
	Population			Population		
Census/Response Mode	Millions	Percent	MWI	Millions	Percent	MWI
Census Total	308.7	100.0	0.56	331.4	100.0	1.48
Self-Response						
Internet	N.A.	N.A.	N.A.	201.6	60.8	0.36
Paper	214.2	69.4	0.21	39.2	11.8	0.28
Telephone	17.5	5.7	0.29	9.2	2.8	0.57
Nonresponse Followup						
Household Respondent	59.7	19.3	0.77	38.9	11.7	3.16
Proxy Respondent	11.9	3.9	5.52	14.1	4.3	17.44
Administrative Records[a]	N.A.	N.A.	N.A.	10.1	3.0	0.22
Other Enumeration	5.3	1.7	0.86	18.4	5.6	0.16

[a] The 2020 row entries for administrative records are smaller than those in Table 3.2 because this table includes administrative records enumerations only, whereas Table 3.2 includes administrative records enumerations plus other types of enumerations in which administrative records were used to assign age (see Box 3.2).

NOTES: MWI, modified Whipple's index; N.A., not applicable. Population values include all ages. MWI values use data for reported ages 23–62 only. The "Other Enumeration" category is a catch-all category, which is not comparable across censuses. Major operations in this category include Group Quarters Enumeration, Military Enumeration, Remote Alaska Enumeration, and Enumeration of Transitory Locations (e.g., camps for people experiencing homelessness).

SOURCE: Calculations by the panel from Census Bureau tabulations from the 2020 Census Edited File. See Disclosure Review Statement; CBDRB-FY23-0180.

consistent peaks at all multiples of 5 from 25–70 years of age for 2020 and from 25–40 years of age for 2010, none of the estimated heaping metrics exceed 0.25.[5] The large swing between ages 17 and 18 for 2020 is due to an abrupt drop going from age 17–18 in the number of enumerations by self-response. Similarly, the age heaping graph for administrative records in 2020 (Figure 3.7) exhibits virtually no evidence of age heaping.

In contrast, there is strong evidence of heaping for responses obtained from NRFU (Figure 3.8). While there are peaks at almost every multiple of five in each census, the 2020 peaks tower over those for 2010, often by a factor of about three-to-one. As might be expected, there are big differences between the results for household and proxy respondents during NRFU. Comparison of the relevant rows of Table 3.3 as well as Figures 3.9 (household member NRFU

[5]Data on self-response by paper in 2020 (not shown) follow the same pattern as self-response by internet—that is, consistent peaks at all multiples of 5 from ages 25–70 and low values of the $\hat{h}_a$ metric.

Copyright National Academy of Sciences. All rights reserved.

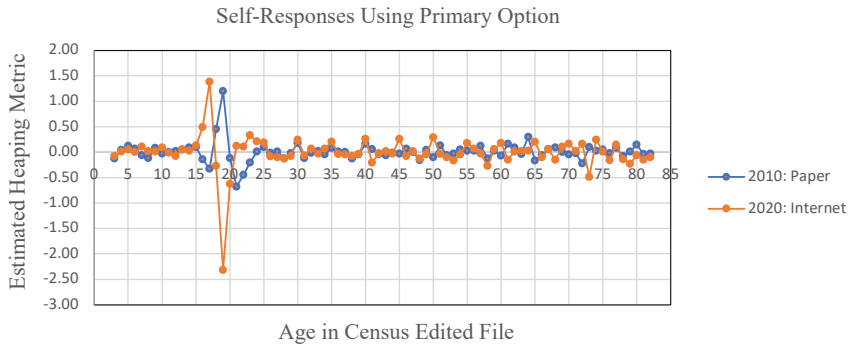


Figure 3.6 Age heaping for respondents who used primary Self-Response option, 2010 Census (paper) and 2020 Census (internet), using the $\hat{h}_a$ metric.

SOURCE: Calculations by the panel from Census Bureau tabulations from the 2020 Census Edited File. See Disclosure Review Statement; CBDRB-FY23-0180.

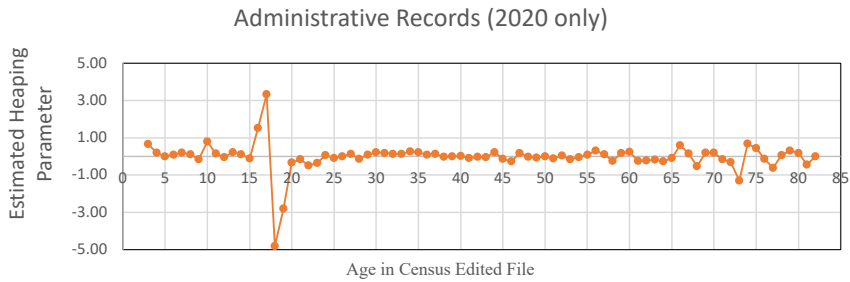


Figure 3.7 Age heaping for respondents obtained by administrative records in the 2020 Census, using the $\hat{h}_a$ metric.

SOURCE: Calculations by the panel from Census Bureau tabulations from the 2020 Census Edited File. See Disclosure Review Statement; CBDRB-FY23-0180.

Copyright National Academy of Sciences. All rights reserved.

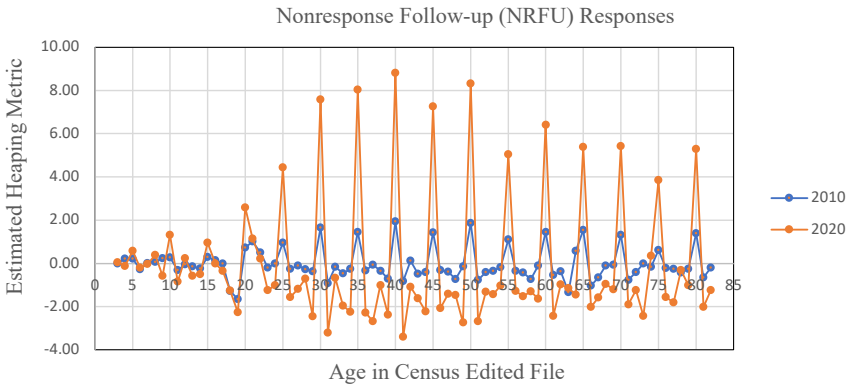


Figure 3.8 Age heaping for respondents enumerated in Nonresponse Followup, 2010 and 2020 Censuses, using the $\hat{h}_a$ metric.

SOURCE: Calculations by the panel from Census Bureau tabulations from the 2020 Census Edited File. See Disclosure Review Statement; CBDRB-FY23-0180.

respondents) and 3.10 (proxy NRFU respondents) indicate that there is at least five times the rate of heaping by proxy respondents as by household members during NRFU.[6] Note, however, the different scales in the various figures: Figure 3.6 (−3 to +2), 3.7 (−5 to +5), 3.8 (−4 to +10), 3.9 (−3 to +5), and 3.10 (−10 to +25).

3.5 AGE HEAPING—CONCLUSION

Conclusion 3.1: Age heaping (excess responses of ages ending in 0 or 5) is a well-recognized indicator of data quality in censuses. Age heaping was much more prevalent in the 2020 Census than in 2010, by an estimated factor of at least 2.5 for ages between 23 and 62 years, for which heaping is most common. This increase was almost completely due to increases in the rate of heaping during Nonresponse Followup (NRFU)—especially from proxy respondents but also from NRFU household respondents. Age heaping was rare in both censuses for self-responses, whether by paper or internet (in 2020 only), and there was no evidence of age heaping associated with administrative records.

[6]Despite constituting only 4% of the full census, data not shown indicate that NRFU proxy responses contributed an estimated 55% of the heaped ages in the 23–62 year range in 2020, while household member NRFU responses contributed another 26%.

Copyright National Academy of Sciences. All rights reserved.

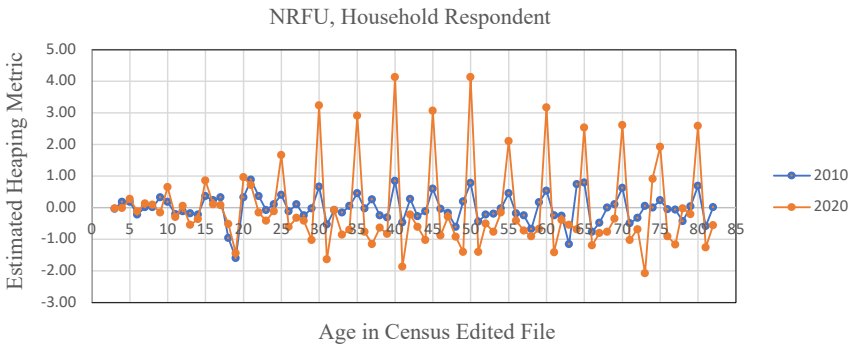


Figure 3.9 Age heaping for Nonresponse Followup resolutions by household member interview, 2010 and 2020 Censuses, using the $\hat{h}_a$ metric.

SOURCE: Calculations by the panel from Census Bureau tabulations from the 2020 Census Edited File. See Disclosure Review Statement; CBDRB-FY23-0180.

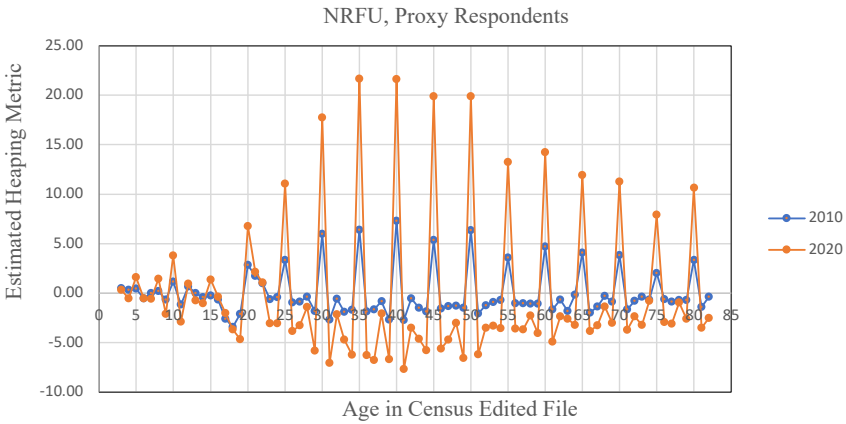


Figure 3.10 Age heaping for Nonresponse Followup resolutions by proxy, 2010 and 2020 Censuses, using the $\hat{h}_a$ metric.

SOURCE: Calculations by the panel from Census Bureau tabulations from the 2020 Census Edited File. See Disclosure Review Statement; CBDRB-FY23-0180.

Copyright National Academy of Sciences. All rights reserved.

The increase in age heaping raises concerns about the accuracy of age data collected in 2020. Some analyses, notably projections of pre-K–12 school enrollment, require accuracy for single-year-of-age distributions for the relevant age range. Other analyses require accuracy for specific ages—for example, numbers of people reaching or projected to reach key retirement ages, such as age 62 for early Social Security benefits or age 65 for Medicare eligibility. Age heaping could bias such analyses substantially.

In our view, data users should be aware of the increased heaping in the 2020 Census and consider either smoothing reported distributions or analyzing data for 5-year bins, for example. Age heaping appears to be rare for children, so it should not greatly affect planning for public schools or calculation of disease and mortality rates for this population. However, other errors may be consequential for these uses, including the large net undercount for young children ages 0–4 (see Figure 4.1 in Chapter 4), high imputation rates for age and date of birth for enumerations obtained in NRFU (see Figure 3.11), and error introduced by the new disclosure avoidance methods for the 2020 Census (see Chapter 11). Analysis by the Census Bureau of the likely combined errors for young children ages 0–4 and older children ages 5–9 and 10–17 would be helpful for important uses of the data.

Looking toward 2030, it would be imprudent to assume that the excessive age heaping in the 2020 NRFU operation, particularly in proxy responses, was due solely to the COVID-19 pandemic and not also to such factors as distrust of government, which may persist or grow worse over time.[7] To better understand and prepare for possible age heaping in the 2030 Census, it would be useful for the Census Bureau to extend the analysis in this chapter. For example, the Census Bureau could calculate age heaping metrics for cross-classifications of age reporting (see Table 3.2) and response mode (see Table 3.3) for substate geographic areas and race and ethnic groups. Similarly, the Census Bureau could stratify the populations in Table 3.4 by the degree of enumeration difficulty (e.g., the Low Response Score in the Planning Database).[8]

[7] According to Pew Research Center (2022), the American public's trust in "the government to do what is right just about always/most of the time" has averaged around 20% since the beginning of the Obama administration (dropping from a prior high of about 45% at the start of the George W. Bush administration). The 2020 Census Barriers, Attitudes, and Motivators survey and focus groups, conducted in 2018, found the chief barrier to census participation was lack of understanding of its purpose which, in turn, was associated with apathy, privacy concerns, fear of repercussions, and general distrust of government. Low levels of education, being young, and being a person of color related to low levels of intent to self-respond. See McGeeney et al. (2019).

[8] See https://www.census.gov/topics/research/guidance/planning-databases.html; the Planning Database is a Census Bureau product that, in previous years, conveniently combines census tract- and block group-level variables from the decennial census (including operational data such as Self-Response rates) as well as deriving some new metrics such as a Low Response Score (previously, a Hard-to-Count Score). The 2021 and 2022 vintages of the Planning Database have not yet included information from the 2020 Census because of the delays in generating 2020 data products.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

3.6 DATA QUALITY IN THE NONRESPONSE FOLLOWUP OPERATION

Perhaps of more importance than the effects of the increase in age heaping on uses of age data from the 2020 Census is what age heaping suggests about the quality of census data for other characteristics and on census coverage itself. Age heaping occurs when a respondent either does not know a subject's exact age or decides it is quicker or easier to give an approximate age. If the reason for age heaping were ignorance of age information, it seems likely that respondent knowledge would also be lacking about other characteristics, including Census Day residence; when age heaping results from a drive to complete the interview as quickly as possible, that could negatively affect all the information collected. It is easy to imagine fraught-with-tension interviewing amidst the COVID-19 pandemic resulting in responses that were as curt and quick as possible.

The evidence indicates that NRFU responses, particularly from proxies, are of poorer quality than self-responses and that their quality was generally worse in 2020 than in 2010. Age heaping, item nonresponse and imputation, and coverage errors tell a similar story. Evidence presented above shows that age heaping is concentrated in NRFU and particularly in proxy responses, and that the extent of age heaping in 2020 was several times that of 2010 and 2000.

Regarding coverage, following a recommendation from the National Research Council (2004a:8–9), the 2010 PES looked at components of coverage error for census operations to help uncover operational factors that might contribute to overcount or undercount of areas and population groups. The 2020 PES followed suit. Table 3.4 shows percentages of correct enumerations, duplicates, other erroneous enumerations, and whole-person imputations for various operations accounting for over 99% of the 2020 Census count. Note that correct enumerations include cases that matched the PES and cases that did not match but were determined to be correct in the census. Also shown is the percentage of total household enumerations contributed by each operation. No estimates are provided for people included in the PES who were omitted from the census, although estimates could be developed for omissions in census households in which at least one person matched a person in a PES household (see Section 4.3).

Table 3.4 shows that NRFU responses obtained from a household member not the head or from a proxy respondent had fewer correct enumerations and more whole-person imputations than all other enumeration modes, including NRFU enumerations from the household head. In fact, NRFU enumerations from the household head were more complete than internet and telephone responses from people who did not have the ID provided to them in Census Bureau mailings.

Regarding item nonresponse, evidence in operational metrics released by the Census Bureau for the nation as a whole and for individual states shows that

Copyright National Academy of Sciences. All rights reserved.

Table 3.4 Percentage of the Household Population by Components of Coverage Error by Census Operation, 2020 Census Post-Enumeration Survey

Coverage Error Component/ Census Operation	% Operation of Enumerations in Households	*Four columns add across to 100% of each operation*			
		Correct Enumerations	Duplicate Enumerations	Other Erroneous Enumerations	Whole-Person Imputations
Census Total	100.0	94.4	1.6	0.6	3.4
Self-Response					
Internet and Telephone ID Response	55.6	96.9	0.9	0.4	3.4
Internet and Telephone Non-ID Response	9.4	91.6	2.3	0.6	5.6
Paper-based Response	11.6	97.1	1.3	0.6	1.6
Nonresponse Followup (NRFU)					
NRFU Head of Household	9.7	94.0	3.3	0.8	1.8
NRFU Other Household Member	2.9	76.6	2.2	0.5	22.7
NRFU Proxy	4.7	87.4	3.7	1.8	7.0
Administrative Records	3.1	94.5	4.3	1.3	—
Update Leave/Update Enumerate	2.0	91.7	4.0	0.6	3.6
Count Imputations	0.6	—	—	—	100.0

NOTES: The estimates in this table are drawn from various tables in the source document; the categories are not necessarily exhaustive or mutually exclusive. In particular, the Update Leave/Update Enumerate category estimates represent those types of enumeration areas—the Update Leave responses consequently overlap with internet, paper, and NRFU responses. All estimates are statistically significantly different from 0 at the 10% level. (Whole-person imputations are a census count and do not have associated sampling error.)

SOURCE: Hill et al. (2022:Appendix Tables 7, 9, 10).

Copyright National Academy of Sciences. All rights reserved.

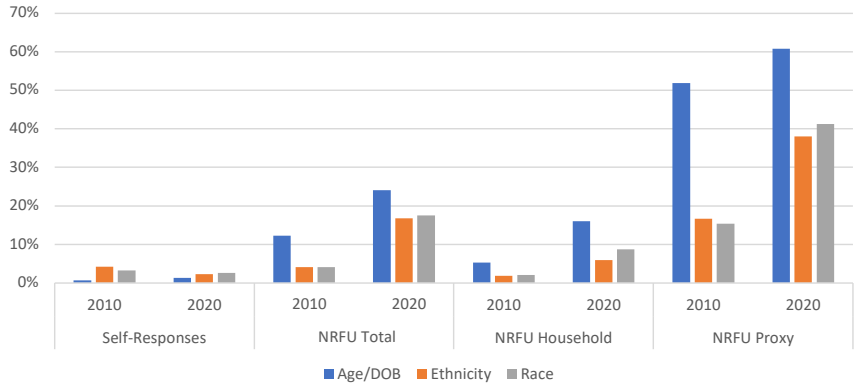


Figure 3.11 Item nonresponse rates for age/date of birth, ethnicity, and race, by type of response, 2010 and 2020 Censuses.

NOTES: DOB, date of birth; NRFU, Nonresponse Followup. Item nonresponse rates refer to missing responses; item imputation rates (see, e.g., Table 10.5) are usually higher because they include cases in which responses had to be blanked and imputed; denominators for percentages are people in occupied units providing the designated type of response (e.g., self-response).

SOURCE: Plotted from 2020 Census Operational Quality Metrics Release 3, Table 2 (U.S. Census Bureau, 2021c).

item nonresponse rates are higher in NRFU, particularly in proxy responses, and that the extent of item nonresponse was higher in 2020 than in 2010 (see Figure 3.11, which shows item nonresponse rates for age/DOB, race, and ethnicity).[9]

> *Conclusion 3.2: Findings about age heaping raise concerns about the overall quality of data collected during 2020 Nonresponse Followup (NRFU), especially from proxy respondents but even from some household members. Other quality indicators, including item nonresponse rates, numbers of people requiring imputation of all characteristics, and coverage errors, reinforce the difficulties of obtaining accurate data in NRFU and indicate that the difficulties increased in 2020 compared with 2010. The greater extent of quality problems in NRFU in 2020 compared with 2010 may be attributable to the COVID-19 pandemic (including the resulting delay in the NRFU schedule), but other factors*

[9] Administrative records used in nonresponse enumeration in 2020 had missing data rates of 3.5% for age/DOB, 28.3% for ethnicity, and 18.1% for race.

Copyright National Academy of Sciences. All rights reserved.

might also be involved, which would need to be addressed in 2030 Census planning.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

– 4 –

Coverage Measurement in the 2020 Census

Every census since 1790 has missed some people who should have been counted, counted some people twice, and counted other people who should not have been counted (e.g., someone born after Census Day). The balance of undercounted and overcounted people determines whether the census experienced a net undercount or overcount for the nation as a whole. In addition, people can be missed in one place and erroneously counted in another place according to census rules regarding people's usual residence. Also, people may misreport information such as age, or may fail to report information such as race or ethnicity, for example, and have race or ethnicity erroneously imputed. These kinds of errors will not affect coverage for the nation but will add to net under or overcount in a geographic area or population group.

Differential undercount—that is, the difference in coverage rates between population groups or geographic areas—is more concerning than the national net overcount or undercount. The reason is that many important uses of census data involve allocating shares of a fixed pie—whether that pie represents seats in the U.S. House of Representatives allotted to the various states, or funds allotted to geographic locales by a formula that uses census or census-derived data for the allocation. Indeed, a census that obtained close to the correct number of people, balancing omissions against erroneous enumerations, but had large differences in net undercount rates between population groups and areas could be less equitable than a census that uniformly missed 1–2% of people across all groups and areas.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

In this chapter, we describe how the U.S. Census Bureau estimates net undercount rates using Demographic Analysis (DA) and a Post-Enumeration Survey (PES) with dual-system estimation (DSE), and how the PES is used to estimate the contribution of errors such as omissions or duplications (termed gross errors) to the net undercount. We then present DA and PES results for the 2020 Census compared with the 2010 Census. The major finding is that, while the total population in both the 2020 and 2010 Censuses closely matched the DA- and PES-estimated totals, the differential net undercount increased from 2010 to 2020. Net *undercounts* increased substantially for Hispanic people, Black people, and American Indians on reservations, while net *overcounts* increased for White and Asian people.[1] We discuss problems affecting the 2020 DA and PES estimates and additional research to evaluate the two methods and expand their use to identify potential ways to reduce the differential net undercount in the 2030 Census. The chapter ends with our conclusions and recommendations.

4.1 ESTIMATING COVERAGE ERRORS

There are several methods to estimate coverage errors in the census (see Box 4.1). In recent censuses, the Census Bureau has used DA and PES exclusively. Each method has strengths and weaknesses.

4.1.1 Demographic Analysis

DA is generally considered as close to a gold standard as is possible for estimates of the population as of April 1 of a census year. The current DA program is based on decades of analysis of censuses and administrative records; since the advent of the Medicare program in 1965, it has used separate methods for two age groups. For people born in 1945 or later, when birth records were complete across the nation, the 2020 DA estimates reflect the addition of births over the decades, the subtraction of deaths, and the addition of estimates of net international migration (from various sources, including the American Community Survey [ACS]). For people ages 75 and over (born in 1945 and earlier), the 2020 estimates are based on Medicare enrollment data, adjusted for people not covered. The Medicare enrollment method is more accurate for this group than using birth records given gaps in records for older people. DA estimates are largely independent of the census.[2]

[1]The 2000 PES also estimated that Black and Hispanic people were undercounted and White and Asian people overcounted (National Research Council, 2004a:Table 6.7, Revision II A.C.E. estimates).

[2]See Jensen et al. (2020). Note that censuses from 1970 through 2010 used Medicare enrollment data for people ages 65 and older (Robinson, 2010).

Copyright National Academy of Sciences. All rights reserved.

Box 4.1 Methods for Estimating Census Coverage

Demographic Analysis—Constructs estimates of the population from administrative records (birth records, death records, immigration-related data, Medicare records).

Post-Enumeration Survey (PES) with Dual-System Estimation (DSE)—Draws a sample of blocks, constructs a comprehensive address list for the blocks, and conducts a survey (P-sample) of residents of those addresses; then matches the P-sample records with the census enumerations in the sample blocks (E-sample) to estimate correct matches, omissions of P-sample people from the E-sample, and erroneous enumerations in the E-sample; uses DSE to estimate the total population (which will include the estimated number of people not captured in either the P-sample or the E-sample). (Note: A PES can be conducted without using DSE by, in effect, taking the census again in a sample of blocks using highly trained interviewers. This method has been shown to fall short of the census count because it is not independent of the census; therefore, it is not appropriate for coverage estimation.)

Reverse Record Check—Draws samples of people counted in the previous census, postcensal births, postcensal immigrants, and people determined through coverage evaluation to have been missed in the previous census; locates the sample people to determine if they are still alive and in the country; estimates the total population.

Administrative Records Matches—Matches records of people in one or more administrative systems (e.g., Social Security records) to the census; sometimes this method is used for specific populations such as people eligible by age for Medicare.

Participant Observation—Participant observers, such as trained cultural anthropologists, generate lists of people in the neighborhoods to which they are assigned, to estimate the total population in those areas. This method is not feasible on a large scale but can be helpful in studying coverage errors in specific populations, such as Asian language communities.

NOTES: To estimate net overcount or undercount, the census population total is compared with the population total from the coverage estimation method, with the latter serving as the denominator for coverage rates. For further discussion of coverage estimation, see National Research Council (1985:Ch. 4), reissued as National Academies of Sciences, Engineering, and Medicine (2015:Ch. 4); National Research Council (2004a:Ch. 5, 6); and National Research Council (2008).

Copyright National Academy of Sciences. All rights reserved.

DA estimates are compared with census counts to produce net undercount estimates; the method cannot be used to produce estimates of gross errors, such as omissions or duplications. DA estimates are limited to the nation because estimates of interstate migration are not sufficiently accurate for the preparation of state estimates that can be compared with census results.[3] DA estimates are produced for single years of age, sex, and Black and non-Black people (separate estimates for Black Alone and Black Alone or in Combination). Estimates are not produced for other race groups because of gaps in birth- and death-record identification of race. For 2020, experimental DA estimates were produced for Hispanic people ages 0–29 (estimates for ages 0–19 were produced for 2010). Beginning in 2010, the Census Bureau has released a range of DA estimates, to reflect uncertainty in several components of the estimation. For 2020, the Census Bureau released three series—low, middle, and high—which used varied assumptions about births, deaths, net international migration, and Medicare enrollment for people ages 75 and older (see Jensen et al., 2020:Tables 3–8).

The Census Bureau released its 2020 DA estimates in December 2020.[4] In March 2022, it released net undercount estimates, comparing the 2020 DA estimates with the 2020 Census population counts, by single years of age and sex.[5] As of this writing, the Census Bureau has not yet released net undercount estimates by race (see Section 4.3). Data are now available from the Demographic and Housing Characteristics (DHC) File to estimate net undercount for Hispanic children and young adults (up to age 29). The range of estimates for these groups between the low and high DA series is so great as not to be useful (see Section 4.3.1 for a discussion of the estimates and the need for further research to narrow the range).

4.1.2 Post-Enumeration Survey with Dual System Estimation

The other method used to evaluate coverage in recent censuses is DSE applied to the results of matching the responses from a PES (P-sample) to census enumerations (E-sample) in a sample of blocks (see Table 4.1). The goal is to estimate the true population (the bottom right cell shown with darker shading) using some of the estimates that can be constructed for other (lightly shaded) cells in the table.

[3]The Census Bureau plans to release experimental 2020 estimates of net overcount or undercount of children ages 0–4 for states and counties.

[4]See https://www.census.gov/newsroom/press-releases/2020/2020-demographic-analysis-estimates.html.

[5]See https://www.census.gov/newsroom/press-releases/2022/2020-census-estimates-of-undercount-and-overcount.html. To permit 2020 Census–DA comparisons by single years of age and sex, when census data were not scheduled to be publicly available until the DHC File was released in May 2023, the Census Bureau prepared a national-level census tabulation for internal use, protected using a differential privacy-based algorithm with a "light touch."

Copyright National Academy of Sciences. All rights reserved.

Table 4.1 Components of the Dual-System Estimator of the True Population

	P-Sample: Correctly Enumerated in PES?		
E-Sample: Correctly Enumerated in Census?	Yes	No	Total
Yes	N_{11} Census-PES Matches	N_{12} In Census, Out of PES	N_{1+} Correct Census Total
No	N_{21} In PES, Out of Census	N_{22} Missing from Both	N_{2+} Missing from Census
Total	N_{+1} Correct PES Total	N_{+2} Missing from PES	N_{++} **True Population Total**

NOTES: PES, Post-Enumeration Survey.

SOURCE: Panel generated.

The DSE estimate of the true total population exploits the assumed (statistical) independence of the census results and the PES results. Based on this assumed independence, the ratio of the number of census-PES matches (N_{11}) to the correct census total (N_{1+}, which is the percentage of people captured by the PES conditional on their capture in the census) should equal the ratio of the correct PES total (N_{+1}) to the unknown true total population (N_{++}, which is the unconditional percentage of people captured by the PES). The resulting expression can be rearranged into Equation 4.1, and the resulting DSE Total estimate can then be used to calculate the net coverage rate, as in Equation 4.2:

$$\text{True (DSE) Total}(N_{++}) = N_{1+} \times \frac{N_{+1}}{N_{11}} \tag{4.1}$$

$$\text{Net Coverage Rate} = 100 \times \frac{\text{Census Total} - \text{DSE Total}}{\text{DSE Total}} \tag{4.2}$$

To support this assumption of independence, the Census Bureau keeps census and PES processes operationally separate. In particular, the Master Address File is used as the frame for the census but *not* the PES, which instead fields a separate Independent Listing operation to construct its address list in sample blocks. Further, the Census Bureau makes extensive use of post-stratification to account for higher or lower rates of capture, in both the census and PES, by factors such as housing tenure (owner or renter) and the race and ethnicity characteristics of the residents. More recently, logistic regression has also been used to essentially condition the above DSE calculation on the values of these factor variables, resulting in more homogeneous subsets of the population with respect to their capture probabilities.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

In addition to the independence assumption, these equations fail to take into consideration various complications that stem from incomplete PES and census data, duplicate enumerations, and erroneous enumerations. Further, the methodology is based on various additional assumptions, especially that the matching of the data from the census and the PES needs to be carried out with few errors. One aspect of incomplete data is that not all census enumerations can have their match status to the PES ascertained because they are not "data-defined," meaning that they did not have sufficient information to be accepted for census processing and so had to be wholly imputed. Given various corrections to deal with duplicates, erroneous enumerations, and imputations, the Census Bureau estimates Coverage Correction Factors (CCFs) as a modification of the DSE estimators. The quality of the CCFs is therefore impacted by errors in the estimation of the number of duplicates, errors in the estimation of the number of erroneous enumerations, errors in the treatment of incomplete data, and the impact of false non-matches and false matches. CCF quality therefore depends on the frequency of such errors, which needs to be assessed to judge the success of the estimates. The resulting complexity of PES data collection and estimation is summarized in Box 4.2, which provides a simplified overview of the major steps.

From the 2020 PES, the Census Bureau released national estimates of net and gross coverage errors in March and May 2022 for people in households by age and sex groups, race and ethnicity, household relationship, housing tenure (i.e., own, rent), and geography (i.e., 4 regions, 50 states, and District of Columbia).[6] The Census Bureau also provided estimates of components of error for various census operations. The 2020 PES estimates generally have 60–80% larger standard errors than in 2010 because the sample size was smaller to start with and the COVID-19 pandemic and other factors led to increased household nonresponse (see Section 4.3).

4.2 2020 COVERAGE RESULTS

4.2.1 Total Population

Table 4.2 provides the total population net undercount rates for the 2020 and 2010 Censuses for the low, middle, and high series of DA estimates. In each year, the census counts for the total population are on target with the DA middle

[6]See the Census Bureau's 2020 PES press kit at https://www.census.gov/newsroom/press-kits/2021/post-enumeration-survey.html. Unlike in 2010 (see Davis and Mulligan, 2012:Attachments 1, 2), the Census Bureau did not release coverage error estimates for large counties or places in 2020, determining that the uncertainty due to modeling error for substate estimates did not warrant their production. It did release national estimates of net and gross coverage errors by several characteristics for housing units, and also released household population and housing unit coverage estimates for Puerto Rico in August 2022, which are not further discussed.

Copyright National Academy of Sciences. All rights reserved.

Box 4.2 Major Steps in Post-Enumeration Survey Data Collection and Estimation

- Draw a sample of blocks and construct a comprehensive address list for them (P-sample) prior to and operationally independent of the census. (In both 2010 and 2020, the Post-Enumeration Survey [PES] excluded group quarters and remote Alaska.)
- Subsample small blocks with 0–2 housing units for efficiency; similarly, subsample housing units in large blocks.
- Conduct a survey of the P-sample addresses after census Nonresponse Followup is well underway (to promote independence of the P-sample), asking household members about their residence as of Census Day.
- Match the P-sample records with the census enumerations (E-sample) in the sample of blocks to estimate P-sample and E-sample matches (N_{11} in Table 4.1), P-sample people missed in the E-sample (N_{21}), and E-sample correct enumerations missed in the P-sample (N_{12}).
- Follow up cases in the field that have undetermined enumerations and match status or both, to try to resolve their status.
- Estimate final match and enumeration statuses.
- Calculate preliminary Coverage Correction Factors (CCFs) for population groups (called domains)[a] on the premise that the assumption of independence of being captured in the P-sample and of being captured in the E-sample holds for such domains. There are preliminary CCFs that estimate the probability that a census enumeration is data defined and the probability that an enumeration is correct, and these are applied to the DSE estimate from Equation 4.1.
- Apply Demographic Analysis sex ratios to the preliminary CCFs for adult men and women by age group (18–29, 30–49, 50–64, and 65 and over) and race (Black Alone or in Combination and all other) to adjust for systematic undercounts of adult men (particularly Black men) and overcounts of adult women in the census. (These adjustments were first used in the revision of the 2000 PES estimates.)
- Sum the final CCFs for a group to estimate its true population, which is then compared with its total census count to obtain the net coverage rate.
- Estimate components of error (gross coverage errors): correct census enumerations, erroneous census enumerations (duplicates and people who should not have been counted, such as babies born after Census Day), whole-person census imputations, PES people who match the census, and PES people omitted from the census.
- Analyze error components for population groups to determine the contribution of, say, omissions and erroneous enumerations, to their net coverage rates, and do the same for census operations (e.g., Nonresponse Followup, Update Enumerate) to determine operations that were particularly prone to such errors as duplications.

SOURCE: Panel generated. See Heim (2022) for a detailed description of the methodology.

[a] CCFs were estimated for groups defined by such variables as tenure (owner/renter), state of residence, type of enumeration area, census tract self-response rate decile, race and Hispanic origin, sex and age, and household relationship.

Copyright National Academy of Sciences. All rights reserved.

Table 4.2 Percentage Net Coverage Error for the Total Resident Population: 2010 and 2020 Demographic Analysis and Post-Enumeration Survey Estimates

	Demographic Analysis			
Year/Series	Low Series	Middle Series	High Series	Post-Enumeration Survey
2010	1.00	0.13	−1.27	0.01
2020	0.22	−0.35	−1.21	−0.24

NOTES: Positive numbers are net overcounts; negative numbers (in red) are net undercounts (the DA or PES estimates are the denominators); DA estimates are for the resident population; PES estimates are for the household population, excluding group quarters and remote Alaska. PES estimates are not statistically significantly different from zero at the 10% level in either the 2010 or 2020 Censuses.

SOURCE: U.S. Census Bureau (2022c).

series estimates within plus or minus 1%. The table also provides estimates of total population net undercount rates for 2020 and 2010 from the PES.

4.2.2 Age and Sex

Figure 4.1 compares 2020 net undercount estimates from the DA middle series for single years of age by sex. Age heaping—that is, net overcounts of ages ending in 0 and 5 and net undercounts of surrounding ages (see Chapter 3)—is evident for both men and women; generally, adult men have higher net undercount rates and lower net overcount rates than adult women—a pattern found in previous censuses. Figure 4.1 also shows substantial net undercounts of children and overcounts of older people, which characterized prior censuses as well. In addition, the 2020 Census substantially overcounted college-age people compared with a modest net overcount of this age group in 2010 (see Figure 3.1).

Figure 4.2 shows net coverage estimates from the 2010 and 2020 PES and the 2020 DA middle series by age and sex groups, with the DA estimates generally considered to be more accurate in the absence of contrary evidence. The three sets of estimates are similar except that:

- The 2020 PES estimated a larger net undercount of young children (ages 0–4) than in 2010, and the 2020 DA estimated an even larger net undercount;
- The 2020 DA estimated a larger net undercount for children ages 5–9 than either the 2010 or 2020 PES; and

Copyright National Academy of Sciences. All rights reserved.

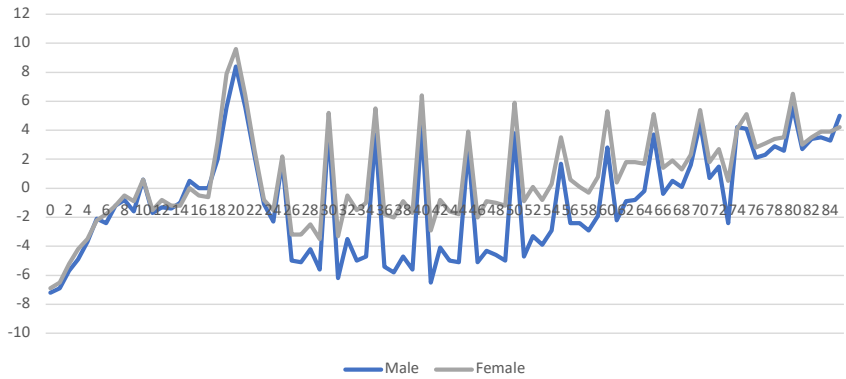


Figure 4.1 Demographic Analysis estimates of percentage net undercount or overcount, by single years of age and sex, 2020, middle series.

SOURCE: Middle series from Table 1—Demographic Analysis Net Coverage Error Estimates by Single Year of Age, Sex, and Series: April 1, 2020 at https://www.census.gov/data/tables/2020/demo/popest/2020-demographic-analysis-tables.html.

- The 2020 DA estimated a net overcount of women ages 18–29 and a negligible overcount of men ages 18–29, while the 2010 and 2020 PES generally estimated net undercounts for these groups. The reason for the discrepant findings for people ages 18–29 is likely because the PES, which excluded college/university student housing along with other group quarters, could not capture the significant problems in the 2020 Census enumeration of college housing (see Chapter 9).

4.2.3 Race, Ethnicity, Housing Tenure, Age and Sex Groups, States

The PES produces estimates for more attributes than does DA, including housing tenure, detailed race and ethnicity, and states, although it produces less detail for age and sex. By housing tenure,[7] the PES estimated that the census had a net undercount of −1.48 percent for renters compared with a net overcount of 0.43 percent for owners in 2020—a similar pattern and magnitude to 2010.

[7]The 2020 PES coverage estimates by housing tenure are statistically different from zero at the 10% level. The PES does not report coverage estimates by household relationship because that variable may be recorded differently between the census and the PES for reasons such as the household having a different reference person completing the two interviews. See Khubba et al. (2022:18).

Copyright National Academy of Sciences. All rights reserved.

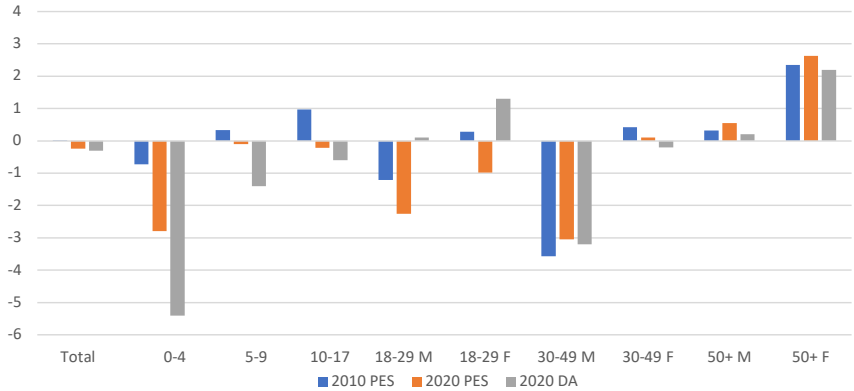


Figure 4.2 Post-Enumeration Survey (2010, 2020) and Demographic Analysis (2020, middle series) estimates of percentage net undercount or overcount, by age and sex groups.

NOTES: DA, Demographic Analysis, covers the resident population; PES, Post-Enumeration Survey, covers the household population excluding group quarters and remote Alaska. All PES estimates are statistically significantly different from zero at the 10% level except estimates for people ages 5–9 (2010, 2020), 10–17 (2020), 18–29 female (2010), and 30–49 female (2020).

SOURCE: Khubba et al. (2022:Table 10).

Figure 4.3 shows estimates of net coverage error by race and ethnicity from the PES for 2010 and 2020. Net overcounts were higher in 2020 compared with 2010 for non-Hispanic White Alone people and substantially higher for Asian people. Net undercounts were higher in 2020 for Black people and American Indian or Alaska Native (AIAN) people on reservations and substantially higher for Hispanic people. Table 4.3 shows how the net differential undercount widened from 2010 to 2020 for Black, AIAN, and Hispanic people versus non-Hispanic White people.

Figure 4.4 shows previously unpublished estimates of net coverage rates from the 2010 and 2020 PES for race and ethnic groups cross-classified by housing tenure (owner/renter) and age/sex groupings. The graphs are for non-Hispanic White Alone people (4.4(a)), Asian people (4.4(b)), Black people (4.4(c)), and Hispanic people (4.4(d)). AIAN people on reservations are not shown because of large standard errors.

Findings that stand out from these detailed cross-classifications include:

- Non-Hispanic White Alone people were almost always overcounted in both censuses, although the percentages are not large except for men and

Copyright National Academy of Sciences. All rights reserved.

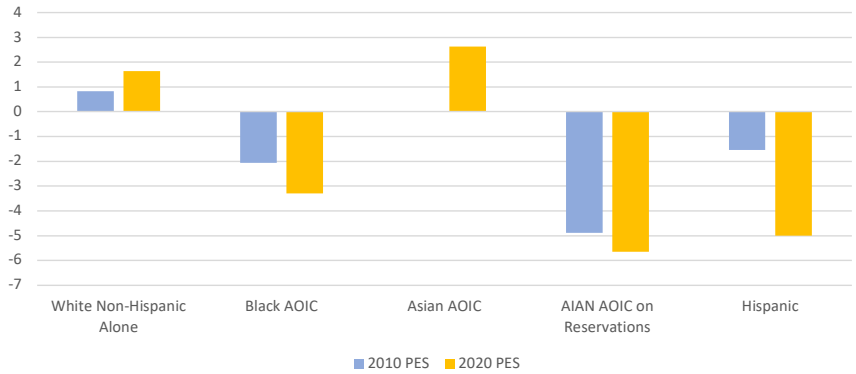


Figure 4.3 Post-Enumeration Survey (2010, 2020) estimates of percentage net undercount or overcount, by race and ethnicity.

NOTES: AIAN, American Indian or Alaska Native; AOIC, Alone or in Combination; PES, Post-Enumeration Survey. Excludes group quarters and Remote Alaska. PES estimates for Native Hawaiian or Other Pacific Islanders and AIAN AOIC in other American Indian areas and living elsewhere not shown because of large standard errors. Estimates for Some Other Race AOIC not shown because they are overwhelmingly Hispanic. All estimates are statistically significantly different from zero at the 10% level except the 2010 estimate for Asian Alone or in Combination (0% net undercount).

SOURCE: Khubba et al. (2022:Table 4).

women ages 50 and over in rental housing, who were overcounted in 2020 by double the rate of 2010.

- Asian people were predominantly overcounted in both censuses but especially so if they lived in rental housing in 2020.
- Black people were generally undercounted in both censuses and particularly so if they lived in rental housing; net undercount rates were somewhat higher for this group in 2020 compared with 2010.
- Hispanic people under age 50 were generally undercounted in both censuses with much higher rates in 2020 than in 2010 and if they lived in rental housing.

Finally, the 2020 PES estimated that 14 states had statistically significant estimates of net over- or undercount, compared with 0 states in 2010. The American Redistricting Project (2022) used the PES estimates for those 14 states to simulate how reapportionment would have turned out in 2020 if the census counts for those states were adjusted for net over- or undercount. It estimated that, compared with the actual reapportionment, California, Illinois, Michigan, New York, Ohio, Pennsylvania, and West Virginia would each still have lost

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Table 4.3 Net Coverage Rates and Differential Net Undercount Rates Comparing White Non-Hispanic People to Specified Groups, 2010 and 2020 Post-Enumeration Survey

Year/Measure/ Group	White Non-Hispanic Alone	Black Alone or in Combination	AIAN Alone or in Combination on Reservation	Hispanic
Net Coverage Rate				
2010 PES	0.83	−2.06	−4.88	−1.54
2020 PES	1.64	−3.30	−5.64	−4.99
Differential Net Undercount (White Non-Hispanic Alone minus specified group)				
2010 PES	—	2.89	5.71	2.37
2020 PES	—	4.94	7.28	6.63

NOTES: AIAN, American Indian or Alaska Native; PES, Post-Enumeration Survey. All net coverage rate estimates are statistically significantly different from zero at the 10% level.

SOURCE: Khubba et al. (2022:Table 4).

a seat; in addition, Minnesota and Rhode Island would each have lost a seat, and Colorado would not have gained a seat. Montana, North Carolina, and Oregon would each still have gained a seat, but Florida and Texas would each have gained three seats instead of one and two, respectively.

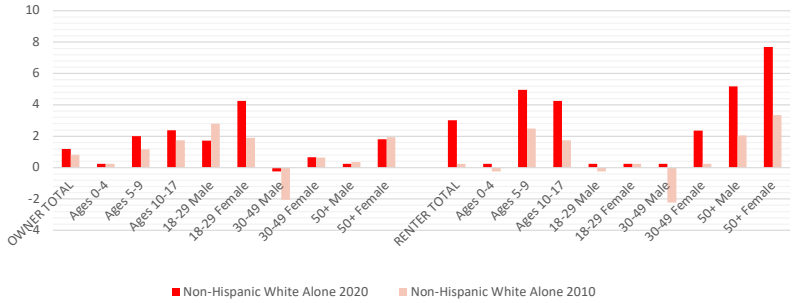
These results should be cited with caution. The method the Census Bureau used to estimate state coverage errors was a "synthetic" method—that is, the Census Bureau did not estimate coverage errors for each state directly but rather as outputs of a national-level logistic regression model that included state as one of the variables in the model (see Heim, 2022:Section 7.2, indicating that the inclusion of a state variable was an improvement over the 2010 PES). Without further evaluation, it is not possible to account for some of the apparently inconsistent results. For example, Florida and Texas both have large Hispanic populations (26% and 39%, respectively), which were severely undercounted nationally as well as in both states. California and New York also have large Hispanic populations (39% and 19%, respectively), yet California was estimated not to have a statistically significant net over- or undercount, and New York was estimated to have a substantial net overcount.

4.2.4 Post-Enumeration Survey—Components of Error

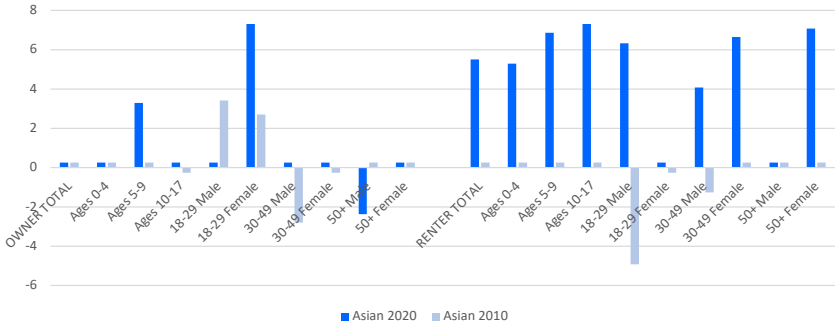
In addition to net coverage error, the PES is used to estimate what are termed gross coverage errors or "components of error." Table 4.4 shows estimates of

Copyright National Academy of Sciences. All rights reserved.

Figure 4.4 Post-Enumeration Survey (2010, 2020) estimates of percentage net undercount or overcount for demographic groups.



(a) Non-Hispanic White Alone people



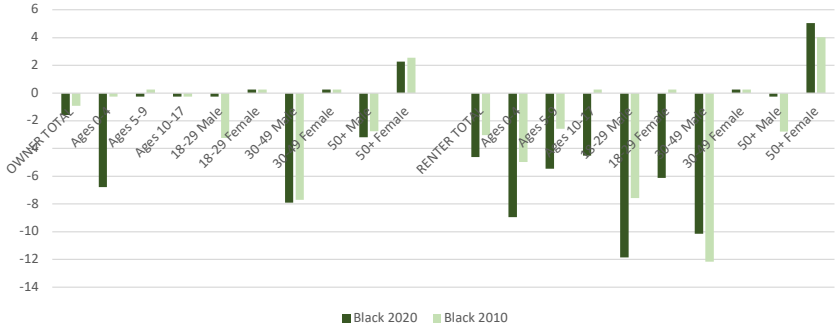
(b) Asian Alone or in Combination people

components of coverage error by race and ethnicity from the PES for 2010 and 2020. On the census side for 2020 (columns 2, 4, 6, and 8), 94.4% of enumerations were correct, 1.6% were duplicates, 0.6% were other types of erroneous enumerations (e.g., counting a baby born after Census Day), and 3.4% were whole-person imputations. On the PES side for 2020 (columns 10 and 12), 94.2% of people in the P-sample matched a census enumeration in the E-sample, while 5.8% of people in the PES were missed in the census. The 2010 PES shows similar results, with a higher rate of duplicates (column 3 relative to column 4) and somewhat lower rates of whole-person imputations (column 7 vs. column 8) and omissions (column 11 vs. column 12).

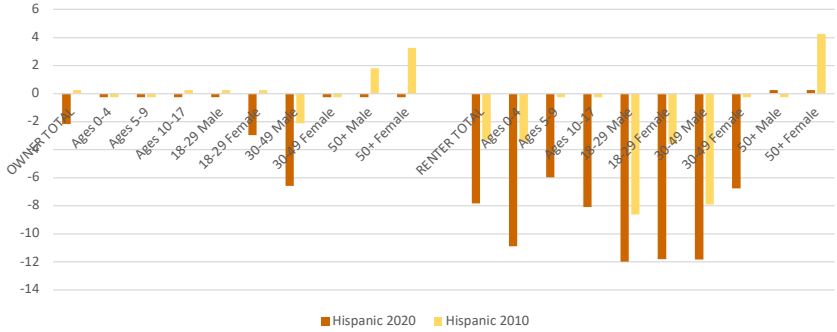
Contributing to the net overcount of Asian people in 2020 (see Figure 4.3) was a lower rate of PES people omitted from the census for this group compared with 2010. Conversely, the net undercount rate for Black people increased in 2020 due to an increase in omitted people. The net undercount rate for Hispanic

Copyright National Academy of Sciences. All rights reserved.

Figure 4.4 (cont.).



(c) Black Alone or in Combination people



(d) Hispanic people

NOTE: Estimates that are very close to zero have values that were set to ±0.25 because they were not statistically significant at the 10% level (the sign of the original estimate was kept).

SOURCES: 2020: Jost and Khubba (2022:Table D). 2010: Keller and Fox (2013:Table D).

Copyright National Academy of Sciences. All rights reserved.

Table 4.4 Percentages of the Household Population by Components of Coverage Error by Race and Ethnicity, 2010 and 2020 Census PES

	Rows sum to 100% of enumerated household population in census (excluding group quarters and Remote Alaska)								*Rows sum to 100% of household population estimated by PES/DSE*			
	Census Correct Enumerations		Census Duplicates		Census Other Erroneous Enumerations		Census Whole-Person Imputations		PES-Matched Enumerations		PES People Omitted from Census	
Year	2010 (1)	2020 (2)	2010 (3)	2020 (4)	2010 (5)	2020 (6)	2010 (7)	2020 (8)	2010 (9)	2020 (10)	2010 (11)	2020 (12)
Total	95.2	94.4	2.7	1.6	0.4	0.6	1.7	3.4	94.7	94.2	5.3	5.8
Race												
White Non-Hispanic Alone	95.4	94.9	2.6	1.5	0.4	0.6	1.6	2.9	96.2	96.5	3.8	3.5
Black AOIC	92.6	92.9	3.6	1.9	0.7	0.7	3.1	4.5	90.7	89.8	9.3	10.2
Asian AOIC	94.7	94.0	2.4	1.8	0.9	0.7	2.1	3.5	94.7	96.5	5.3	3.5
AIAN AOIC on Reservations	90.8	91.7	4.7	4.6	0.4	0.5	4.1	3.2	86.3	86.5	13.7	13.5
Hispanic	93.7	94.2	3.2	1.4	0.7	0.6	2.4	3.8	92.3	89.5	7.7	10.5

NOTES: AIAN, American Indian or Alaska Native; AOIC, Alone or in Combination; DSE, dual-system estimation; PES, Post-Enumeration Survey. Excludes group quarters and Remote Alaska. All estimates are statistically significantly different from zero at the 10% level.

SOURCES: 2010: Mule (2012:Table 9). 2020: Khubba et al. (2022:Table 5).

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

people increased markedly in 2020 due to a marked increase in omitted people, from 7.7% in 2010 to 10.5% in 2020.

Table 4.5 shows estimates of components of coverage error for census operations; it is a direct repeat of Table 3.4 and reprinted here for ease of reference. Clearly, self-responses via paper, the internet, or telephone were the highest-quality enumerations in terms of having the highest percentages correct in the census and among the lowest percentages of duplications, other erroneous enumerations, and whole-person imputations. This association only applies, however, to ID self-responses (i.e., those using the identification code corresponding to a particular address that was provided in the mailings asking households to respond). Non-ID self-responses (9.7% of the total) had a higher-than-average rate of whole-person imputations and a lower-than-average rate of correct census enumerations. Update Enumerate and Update Leave responses were about on par with non-ID self-responses. Administrative records enumerations and Nonresponse Followup (NRFU) enumerations conducted with a household head fell between ID self-responses and Non-ID self-responses in quality. The worst-quality enumerations, with the lowest percentages of correct census enumerations and highest percentages of whole-person imputations, were those obtained in NRFU by either a proxy (nonhousehold member such as a neighbor or landlord) or a household member who was not the head.

While comparisons of 2020 to 2010 would be informative, it is difficult to find categories for which comparable PES estimates were produced for both censuses. Reasonably comparable sets of estimates for 2020 and 2010 pertain to the time when NRFU enumerations were completed. In 2010, there is a definite degradation of quality from early NRFU enumerations to later enumerations. This pattern is not evident for 2020.[8] In both censuses, NRFU enumerations with a household head (2010) or a household member (2020) were of higher quality than NRFU enumerations with a proxy such as a neighbor or landlord. In particular, NRFU household enumerations were correct 93% of the time in 2010 and 94% of the time in 2020, whereas NRFU proxy enumerations were correct only 70% of the time in 2010 and 87% of the time in 2020.[9]

Confirming that self-response is higher quality than other response modes, Figure 4.5 shows rates of census erroneous enumerations, whole-person imputations, and PES people omitted from the census for deciles of census tract self-response rates for 2020.[10] The 2020 Census exhibits the same phenomenon observed in 2000—that quality deteriorates the lower the self-response rate,

[8]For 2010, see Mule (2012:Table 19); for 2020, see Hill et al. (2022:Appendix Table 8).

[9]For 2010, see Mule (2012:Table 21); for 2020, see Hill et al. (2022:Appendix Table 9).

[10]This analysis can include PES people omitted from the census in addition to erroneous census enumerations and whole-person imputations because the variable being examined is the census tract self-response rate.

Copyright National Academy of Sciences. All rights reserved.

Table 4.5 Percentage of the Household Population by Components of Coverage Error by Census Operation, 2020 Census Post-Enumeration Survey

Coverage Error Component/ Census Operation	% Operation of Enumerations in Households	*Four columns add across to 100% of each operation*			
		Correct Enumerations	Duplicate Enumerations	Other Erroneous Enumerations	Whole-Person Imputations
Census Total	100.0	94.4	1.6	0.6	3.4
Self-Response					
Internet and Telephone ID Response	55.6	96.9	0.9	0.4	3.4
Internet and Telephone Non-ID Response	9.4	91.6	2.3	0.6	5.6
Paper-based Response	11.6	97.1	1.3	0.6	1.6
Nonresponse Followup (NRFU)					
NRFU Head of Household	9.7	94.0	3.3	0.8	1.8
NRFU Other Household Member	2.9	76.6	2.2	0.5	22.7
NRFU Proxy	4.7	87.4	3.7	1.8	7.0
Administrative Records	3.1	94.5	4.3	1.3	—
Update Leave/Update Enumerate	2.0	91.7	4.0	0.6	3.6
Count Imputations	0.6	—	—	—	100.0

NOTES: Table repeats Table 3.4 in its entirety and is reprinted here as aid to the reader. The estimates in this table are drawn from various tables in the source document; the categories are not necessarily exhaustive or mutually exclusive. In particular, the Update Leave/Update Enumerate category estimates represent those types of enumeration areas—the Update Leave responses consequently overlap with internet, paper, and NRFU responses. All estimates are statistically significantly different from zero at the 10% level. (Whole-person imputations are a census count and do not have associated sampling error.)

SOURCE: Hill et al. (2022:Appendix Tables 7, 9, 10).

Copyright National Academy of Sciences. All rights reserved.

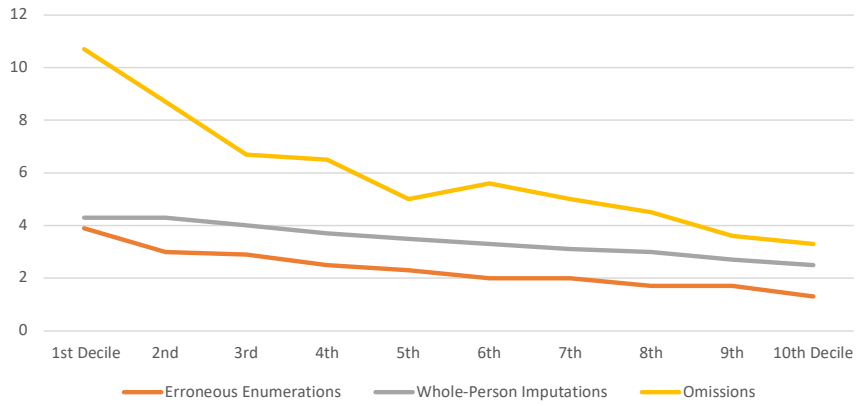


Figure 4.5 Percentages of erroneous census enumerations, whole-person imputations, and Post-Enumeration Survey people omitted from the census for the household population by self-response rate deciles for census tracts, 2020 Post-Enumeration Survey.

NOTES: Each decile contains 10% of the nation's census tracts, ordered by self-response rate from lowest (1st decile) to highest (10th decile). Erroneous enumerations combine duplicate and other erroneous census enumerations; Update Enumerate areas, by design, are enumerated in person and hence fall into the first (lowest) decile of census tract self-response rates. All estimates are statistically significantly different from zero at the 10% level. (Whole-person imputations are a census count and do not have associated sampling error.)

SOURCE: Hill et al. (2022:Table 6).

particularly in terms of omissions of PES people from the census.[11] This phenomenon means that population groups and areas that have lower self-response, which includes census tracts with predominantly Black or Hispanic residents, are at greater risk of net undercount compared with other groups (see Chapter 6).

4.3 CAUTIONS AND AREAS FOR RESEARCH

4.3.1 Uncertainty in Demographic Analysis

Although traditionally viewed as close to a gold standard for estimating census coverage error, DA estimates are not error free. There is sampling error for components derived from the ACS (estimates of foreign-born immigrants

[11]For 2020, see Jost and Khubba (2022:Table D); for 2000, see National Research Council (2004a:Table D.2). A comparable analysis was not produced from the 2010 PES.

Copyright National Academy of Sciences. All rights reserved.

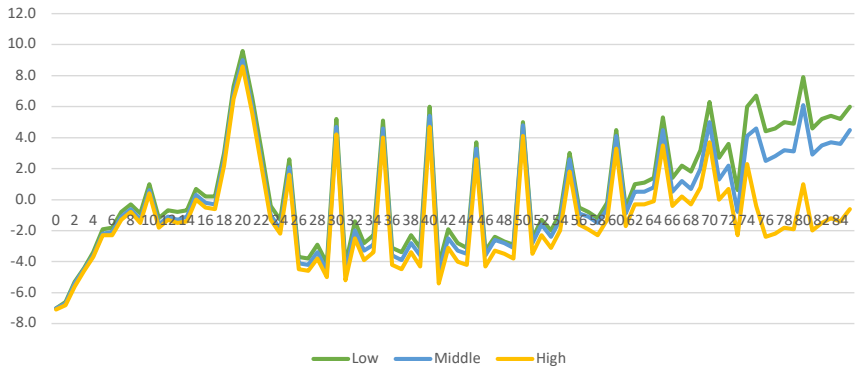


Figure 4.6 Demographic Analysis estimates of percentage net undercount or overcount, by single years of age, 2020, low, middle, high series.

SOURCE: U.S. Census Bureau (2022d:Table 1).

and adjustment of Medicare enrollment for people ages 75 and over in the middle and high series); nonsampling error in birth records prior to 1980 because of registration incompleteness; and error in the estimation process. To endeavor to bound the error in DA estimates, beginning in 2010 the Census Bureau carried out a sensitivity analysis of the DA estimates by constructing lower and higher series around the middle series. In 2010, five series were originally produced, and the middle series was subsequently revised; in 2020, three series were produced. Each series varies on a few key parameters.

For 2020, the components that varied among the low, middle, and high scenarios were births (from 288.4 million in the low series to 289.4 million in the high series); international migration (from 43.2 million to 45.6 million); and Medicare-based estimates (from 21.5 million to 23 million). Death estimates did not vary.[12] Figure 4.6 shows net overcount and undercount estimates from the three series for 2020 by single years of age. The series vary by less than a percentage point through age 19, after which there is more divergence. Around age 75, the high series diverges substantially from the middle series (by about five percentage points on average), which, in turn, diverges from the low series (by about two percentage points on average). In fact, the high series at these older ages produces as many net undercounts as overcounts, whereas the middle

[12]Jensen et al. (2020:Table 3) (see also Table 2, which describes all components of the DA estimates, and Table 8, which describes the assumptions used for the various components for the low, middle, and high series).

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

and low series are always in net overcount territory. The Medicare component is responsible—the difference of about 1.5 million between the low and high series, which affects only people ages 75 and over, is almost the same as the difference in the DA estimates for this group.[13]

For 2020, for DA estimates for people ages 75 and older, the Medicare component equals actual Medicare enrollment in the low series. For the middle series, Medicare enrollment was adjusted upward using unenrolled estimates from 5-year ACS data for 2015–2019 (estimating that 1.7% of people ages 75 and over were not enrolled). For the high series, the middle series-adjusted Medicare estimates were inflated by an arbitrary 5% (estimating that 6.8% of people ages 75 and over were not enrolled). In 2010, Medicare data were used for the entire population ages 65 and older. The adjustment factors, taken from the Current Population Survey, were smaller than the ACS adjustment factors in 2020, and the differences between the high and low series affected ages between 65–70 but not higher ages.[14]

Figure 4.7 shows net overcount and undercount estimates from the three DA series for 2020 for Hispanic and non-Hispanic people for four age groups: 0–4, 5–9, 10–17, and 18–29. The DA high and low series for Hispanic children ages 0–4 differ by 14 percentage points (from a slight net overcount to a 14% net undercount), compared to a much narrower range for non-Hispanic children ages 0–4 of 5 percentage points (from a net undercount of 7% to a net undercount of 2%). Similarly, the ranges for older Hispanic children and young adults are 14 percentage points (ages 5–9), 13 percentage points (ages 10–17), and 12 percentage points (ages 18–29), compared with 4 percentage points (ages 5–9 and 10–17) and 2 percentage points (ages 18–29) for older non-Hispanic children and young adults. The DA low series coverage estimates imply unrealistically that coverage for Hispanic children ages 0–4 was on target while young non-Hispanic children had a large net undercount. Conversely, the low series estimates imply that Hispanic young adults ages 18–29 had a large net overcount while coverage of non-Hispanic young adults was on target. These results conflict with what is known about coverage errors for Hispanic and non-Hispanic people from the PES (see Figure 4.4, particularly panel (d)).

Examining the assumptions built into the 2020 high and low DA series for Hispanic children and young adults, it appears that differences in how Hispanic births were determined are a primary reason for the differences in estimates, especially at younger ages. For the low series, Hispanic births were assigned by the ethnicity of the mother, while for the high series, babies were classified as Hispanic if either parent was reported to be Hispanic, resulting in a larger number of Hispanic births (see Jensen et al., 2020:Tables 5–8).[15] The results

[13]Compare Jensen et al. (2020:Table 3) with U.S. Census Bureau (2022d:Table 1).

[14]Devine et al. (2012).

[15]The Census Bureau used a process that links census records for children and parents—KidLink—to try to assign race and ethnicity as they would be reported in a census or survey, instead

Copyright National Academy of Sciences. All rights reserved.

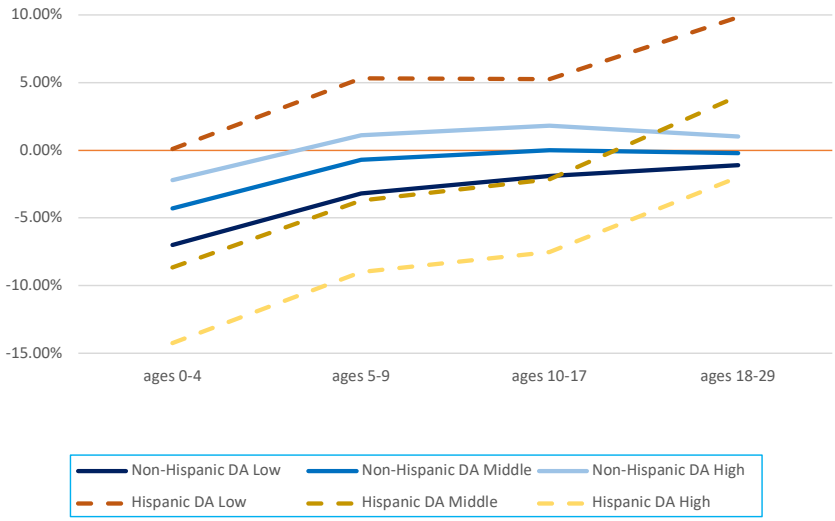


Figure 4.7 Demographic Analysis estimates of percentage net undercount or overcount, by ethnicity, children and young adults, 2020, low, middle, high series.

SOURCE: Calculations by the panel from tables available at data.census.gov for the 2020 Census: Table PCT12 (total), PCT12H (Hispanic); and in U.S. Census Bureau (2022d:Table 3 (total), 1C (Hispanic)) for the DA low, middle, and high series. Note that the non-Hispanic estimates are the result of subtraction of the Hispanic estimates from the total estimates.

are the opposite for non-Hispanic births because they were calculated as a residual. Consequently, the low DA series shows more complete coverage—and even overcounts—of Hispanic compared with non-Hispanic children and young adults and vice versa for the high series. Further research on these large differences is clearly needed.

There are also differences in estimates of net international migration comparing the low to high DA series for Hispanic children and young adults, although they are much smaller than the differences in estimates of births (see Jensen et al., 2020:Table 5). Although Jensen et al. (2020) do not break out the contributions of births and net international migration for Hispanic people by age, it is probable that assumptions about net international migration play a bigger role in differences between the low to high DA series for older age groups. The Census Bureau plans to release an experimental set of DA estimates

of being assigned by other people (e.g., hospital personnel) as on birth records, to classify births for the middle DA series (see Jensen and Kennel, 2022:4–6).

Copyright National Academy of Sciences. All rights reserved.

for Hispanic people ages 30–39, and it will be important to assess the role of various assumptions for net international migration in these estimates and to conduct research to try to narrow the range of estimates.

4.3.2 Delays in Demographic Analysis Net Undercount Estimates by Race

No comparisons of DA estimates for Black and non-Black people to the 2020 Census results have been released to date. To permit race comparisons, census race data must be modified—specifically, people who checked "Some Other Race" (SOR) and neither checked another category nor wrote in any specific race must have a specific race imputed for them because the DA estimates derive from vital records and other sources that historically have not always had an SOR category. The Census Bureau therefore creates a "Modified Race File" to support DA as well as its annual population estimates for the nation, states, and counties. Compared with 2010, the Census Bureau experienced delays in creating a 2020 Modified Race File because it had not developed appropriate confidentiality-protection routines consistent with its new approach to information disclosure avoidance.[16] The Census Bureau hoped to release the 2020 file in summer 2023[17] but, as of this writing, has not produced a file with acceptable accuracy[18] (see Chapter 10 for more information on the Modified Race File and Chapter 11 for information on confidentiality protection in the 2020 Census).

For the Modified Race File, the Census Bureau distributes the SOR Alone category of responses to specific race groups by age and sex. The SOR Alone reassignment method essentially searches for a donor—someone in the household or neighboring area who reported a specific race(s) and Hispanic origin (yes/no) is chosen to represent the person who checked SOR, and the race(s) of the donor are assigned to that person.[19] For people classified as SOR and one or more specific races (i.e., Two or More Races), the Census Bureau assigns a specific race category and removes the person from the Two or More Races category.[20]

[16]See Modified Race Data 2010 at https://www.census.gov/data/datasets/2010/demo/popest/modified-race-data-2010.html, released July 2012.

[17]Private communication from Eric Jensen, U.S. Census Bureau, April 7, 2023. The Census Bureau did not, in any case, want to release the 2020 Modified Race File prior to the May 2023 release of the Demographic and Housing Characteristics file, which contains detailed age, sex, race (including SOR), and ethnicity data from the 2020 Census (see Chapter 11).

[18]See slide 23 in "Briefing on the Base Evaluation and Research Team," available at https://www2.census.gov/about/partners/cac/sac/meetings/2023-09/presentation-briefing-base-evaluation-research-term.pdf.

[19]See U.S. Census Bureau, Population Division (2012) for the 2010 methodology, which is being used for 2020.

[20]In 2020, reassigning people categorized as SOR and a specific race to the specific race will likely create a larger White population, given the new write-in space for the White checkbox and

Copyright National Academy of Sciences. All rights reserved.

4.3.3 Lack of Subnational Demographic Analysis Estimates

Historically, Demographic Analysis has been unable to provide subnational estimates because of the difficulties of estimating not only net international immigration but also internal migration. The Census Bureau plans to produce subnational coverage estimates for children ages 0–4 on an experimental basis, as that is the age group for which estimates are dominated by births and migration should be least problematic. It could be worthwhile to develop subnational estimates for children ages 5–9 as well. The need for subnational estimates is clear from the 2020 PES results, which estimated higher net overcounts and net undercounts than in 2010 for various population groups, which are not distributed equally across the nation.

4.3.4 Uncertainty in Post-Enumeration Survey Results

The previously described complexity of PES data collection and estimation means that there are multiple sources of error in the estimates of each component, such as matched E-sample and P-sample responses, and even more so in estimates of net coverage. Sources of error include:

- Sampling error from sampling census blocks (and subsampling housing units in large blocks) for the P-sample;
- Nonresponse error to the extent that nonrespondents to the P-sample differed from respondents;
- Coverage error in the P-sample, which could miss people and duplicate or otherwise erroneously include people;
- Imputation error from imputing values for missing characteristics for P-sample respondents (imputation error also affects the E-sample);
- Recall error by P-sample respondents as to who lived in the household on Census Day and where they themselves were living on Census Day, April 1 (household members at the time of the P-sample interview could have lived at the P-sample household on Census Day or have moved into the P-sample household from some other household, and some household members on Census Day could have moved out before the P-sample interview);
- Error in estimating match status and enumeration status, including for "unresolved" cases, when even follow-up in the field could not obtain sufficient information to permit an unambiguous classification; and
- Failure of the independence assumption.

Marra and Kennel (2022) provide a comprehensive discussion of potential errors in the 2020 PES but present a method to quantify only the error due to

the coding that assigned people who only checked White but wrote in an Hispanic ethnicity (or other specific race) to White and SOR (or White and other specific race) (see Chapter 10).

Copyright National Academy of Sciences. All rights reserved.

sampling. This means that net coverage rates for, say, race and ethnic groups that are deemed to be statistically different from zero at the 90% confidence level may in fact not be distinguishable from zero were other sources of error taken into account. To date, the Census Bureau has not conducted sensitivity analyses for the PES to attempt to bound errors from the matching process or other sources of error, and no such analyses were conducted for the 2010 PES. The 2000 PES program did include sensitivity analyses, which helped identify components of the data-collection and estimation process that, if somewhat different assumptions were made, could have had significant effects on estimates of net undercount rates. The 2000 PES also included evaluations of several processes (e.g., matching) that included field and clerical work.[21]

4.3.5 Comparison of Errors in the 2010 and 2020 PES

The 2020 PES was conducted under difficult circumstances. The COVID-19 pandemic delayed NRFU in the census and therefore delayed fielding the P-sample interview (the PES independent address listing was completed before the nationwide pandemic-related shutdown in March 2020). The 2020 PES also experienced high levels of nonresponse, likely due to a combination of factors such as fear of contact during the shutdown and heightened distrust of the government. In fact, the Census Bureau extended the P-sample interviewing for an additional period to bring up response rates.

Table 4.6 compares the 2020 and 2010 PES field schedules for the P-sample, while Table 4.7 compares features of the 2010 and 2020 PES that could introduce error. While not explicitly called out in Table 4.7, the delayed schedule for the P-sample interviewing in 2020 (Table 4.6) likely impaired the accuracy of responses about Census Day residence because of the longer recall period. Overall, it is evident that the quality of the 2020 PES was less than that of the 2010 PES in many respects, underscoring the importance of additional analysis of 2020 PES quality, including sensitivity analysis.

4.3.6 Limited Analysis of Components of Coverage Error

A limitation of the 2020 PES program thus far is that the assessments of components of error for census operations are not always comparable with the assessments of 2010, even when comparability could be achieved (see preceding

[21]See National Research Council (2004a:Ch. 6). The original 2010 PES coverage estimation results, which estimated a net undercount, were inconsistent with the DA results, which estimated a slight net overcount. This finding and concerns about some aspects of the 2010 PES led to additional evaluations involving original data collection and analysis. The results of this work demonstrated that the 2010 PES classified many duplicate enumerations in the E-sample as correct and missed in the P-sample. Reanalysis led to a lower census correct enumeration rate, a higher match rate, and a lower estimate of net undercount. See also Mulry and Spencer (1991) for the development of a total error model for the 1990 Census based on the 1988 test census.

Copyright National Academy of Sciences. All rights reserved.

Table 4.6 Schedule of Person Interviewing Field Operations, 2010 and 2020 Post-Enumeration Surveys

Milestone	2010 PES	2020 PES
P-Sample Person Interviewing	8/15/10–10/15/10	9/23/20–12/18/20
P-Sample Second Round of Person Interviewing (to boost response rates)	— (not needed)	2/11/21 – 3/20/21
Field Follow-up of Unresolved Person Cases	1/2011–3/2011	6/14/21–9/17/21

NOTES: PES, Post-Enumeration Survey.

SOURCE: 2010: Whitford (2008). 2020: Marra and Kennel (2022:Table 1).

section). It would be useful, for example, to estimate, for 2010, the components of error for NRFU interviews with a household member other than the head (as in 2020) and for 2020 to estimate the components of error separately for Update/Enumerate and Update/Leave operations (as in 2010).[22] Further, for 2020, it would be useful to produce estimates of components of error for exhaustive, mutually exclusive categories of enumerations (e.g., self-responses versus NRFU responses in Update Leave areas).

It would also be useful to conduct analyses like the one included in National Research Council (2004a:Table D.1). That analysis examined not only duplicates and other erroneous enumerations in the 2000 E-sample, but also omissions of P-sample people from E-sample households, for E-sample households classified by owner/renter and mail/enumerator return status. This type of analysis is limited to households with at least one member who matches between the E-sample and P-sample but could be illuminating nonetheless, particularly if conducted for additional types of census operations (e.g., proxy or administrative records NRFU responses). Given that increases in omissions appeared to drive higher net undercount rates in 2020 for groups such as Hispanic people, an analysis of E-sample and P-sample households with at least one common member seems imperative to understand sources of and possible ways to reduce undercount in 2030.

Additionally, E-sample and P-sample households could be partitioned into two groups—those addresses, households, or individuals for which attempts at enumeration resulted in a census component error (e.g., an omission) and those that did not. Analyzing the two groups using discriminant analysis or

[22]In the 2010 Census, the Census Bureau used the slash (/) character in characterizing these two operations, to distinguish the address verification component of the operation ("Update") from the questionnaire/interview component ("Leave" or "Enumerate"), but no slash is used in the 2020 Census operation names.

Copyright National Academy of Sciences. All rights reserved.

Table 4.7 2010 and 2020 Post-Enumeration Survey Features Relevant for Error

Feature	2010 PES	2020 PES
1—P-Sample Size (addresses)	171,000	161,000
2—Interviews	140,000	114,000
3—Noninterviews	5,300	23,000
4—Vacant or Nonexistent	26,800	24,200
5—Noninterview Rate (row 3/(row 2 + row 3))	3.6%	16.8%
6—P-Sample Size (listed people)	393,000	345,000
7—P-Sample People with Insufficient Information for DSE	13,000	51,000
8—P-Sample People after Whole Households with Insufficient Information Treated as Noninterviews	383,000	301,000
9—Remaining P-Sample People with Insufficient Information (requires imputation of P-sample inclusion and match statuses)	6,400	12,500
10—P-Sample Insufficient Information Rate (row 9/row 8)	1.7%	4.2%
11—E-Sample People	384,000	397,000
12—E-Sample People with Insufficient Information (requires imputation of enumeration and match statuses)	13,000	40,000
13—E-Sample Insufficient Information Rate (row 12/row 11)	3.4%	10.1%
14—P-Sample People with at Least One Characteristic Imputed (%)	6.6%	15.7%
15—E-Sample People with at Least One Characteristic Imputed (%)	15.0%	16.8%
16—Unresolved Match Status Rate (P-Sample)	1.9%	5.6%
17—Unresolved Enumeration Status Rate (E-Sample)	4.8%	11.6%

NOTES: PES, Post-Enumeration Suvey; DSE, dual-system estimation.

SOURCE: Beaghen et al. (2022:Tables 1, 2, 3, 8, 11); Phan and Lawrence (2022:Table 7).

Copyright National Academy of Sciences. All rights reserved.

classification trees with demographic, neighborhood, and other variables as predictors could help assess the quality of each census process in enumerating the population. In turn, the results could provide a basis for considering specific processes for future improvements. See Appendix D.1 for an outline of such an approach.

Having made this suggestion, it must be pointed out that such regression analyses or discriminant analyses, and many of the analyses in this report, compare groups of households that potentially differ for reasons other than their mode of census response or degree of PES error. Basing decisions on such analyses risks assuming correlations and related statistics are causal inferences, when they are associative.

4.3.7 Limited Granularity of Coverage Estimates

Finally, the inability to provide granular geographic or population group coverage estimates is a drawback of the PES (although the PES provides considerably more information than DA in this regard). At the subnational level, direct coverage estimates from the PES have high variability—even if they are unbiased, the variability is too great for usable assessments of population group coverage by state or coverage for substate areas, for example. Careful regression or machine learning models, however, could produce focused coverage estimates with sufficiently low variance and bias (i.e., low mean square error) to support such assessments. For example, estimated coverage for Hispanic people could combine information to adjust for urban/rural, region, and other relevant factors. Models would need to be weighted by estimated variances and possibly by survey weights. The proposed approach would be similar to that used for decades in the Census Bureau's Small Area Income and Poverty Estimates program.

Another approach that could shed light on reasons why some groups are missed in the census more than others and other groups double counted is to conduct a large matching study for 2020 Census enumerations in a sample of census tracts using data from the ACS for 2019–2021 and administrative records from a wide variety of sources—not just those used for enumeration in 2020. The goal would not be to generate another set of coverage estimates for 2020 but to learn as much possible about which groups of people show up in which kinds of data sources to generate testable hypotheses for methods to improve coverage in 2030, particularly for undercounted race and ethnic groups. The study could also examine the suitability of administrative records to improve the count of certain populations, such as young children and Medicare-age adults.

The advantage of using the ACS for matching is that it provides a much larger sample than the 2020 PES—a combined sample for 2019–2021 would total over 5 million households. It also includes group quarters, which the PES did not. ACS interviews farther away from Census Day would be harder to

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

match because people may have moved or changed household composition in the interim. For analysis purposes, ACS interviews could perhaps be inversely weighted according to their distance in time from the census.

To hone in on duplicate enumerations in the census involving more than one address, the Census Bureau could match the set of 2020 Census records against itself. This would identify situations such as those involving people who move seasonally between homes in colder and warmer parts of the United States ("snowbird" and "sunbird" populations), other people with two or more homes, college students enumerated both at college and at home, children enumerated by both parents in joint custody arrangements, and households in apartment buildings enumerated more than once due to apartment number mix-ups. The goal would be to provide an evidence base for methods to reduce duplications involving more than one address in the 2030 Census. An added goal could be to develop a set of well-tested decision rules that would allow the Census Bureau to remove such duplicates, just as it has, for several censuses, used an algorithm to drop duplicate enumerations within a household.

4.4 CONCLUSIONS AND RECOMMENDATIONS

Conclusion 4.1: Comparing the 2010 and 2020 Census net population coverage estimates from Demographic Analysis (DA):

- *For total population, both censuses, based on the middle series of DA estimates, closely approximated the estimated population total (census erroneous enumerations almost offset census omissions).*
- *For age and sex, the 2020 Census undercounted children ages 0–4 more than the 2010 Census. The 2020 Census also overcounted college-age people substantially more than the 2010 Census and, similar to the 2010 Census, overcounted people ages 50 and older.*
- *To date, no estimates of net coverage error for the 2020 Census by race are available because production of the necessary census race data was delayed by development of confidentiality-protection routines.*

Conclusion 4.2: Comparing the 2010 and 2020 Census population coverage estimates from the Post-Enumeration Survey (PES):

- *Differences among race and ethnic groups widened substantially in 2020, with adverse implications for use of the data to allocate fixed resources, such as representation, funding, and services. Some groups saw increases in net overcounts, particularly non-Hispanic White Alone people and Asian people, and other groups saw increases in net undercounts, particularly Black people,*

Copyright National Academy of Sciences. All rights reserved.

Hispanic people, and American Indians and Alaska Natives on reservations.

- *Increases in net undercounts in 2020 for groups such as Black and Hispanic people were driven by increases in omissions from the census (as measured by a lower match rate for people in the P-sample to the E-sample in the PES). Omissions, in turn, were higher in census tracts with lower rates of self-response.*
- *Net overcounts and undercounts were more pronounced for people of various groups in rental compared with owner housing, particularly in 2020 compared with 2010.*

Conclusion 4.3: With respect to the quality of the 2020 Post-Enumeration Survey (PES) and Demographic Analysis (DA) estimates:

- *The quality of the 2020 PES deteriorated compared with 2010, due to smaller sample sizes, higher nonresponse rates, higher missing data rates, higher rates of unresolved match status and unresolved enumeration status, and delays caused by the COVID-19 pandemic.*
- *While DA continued to set a standard for estimation of the total population in 2020, it showed a wide range of net coverage estimates, based on Medicare records and assumptions about completeness of coverage, for people ages 75 and older among the low, middle, and high series of estimates. The 2020 DA coverage estimates series for Hispanic and non-Hispanic children and young adults also showed a wide range, largely due to differences in metnods for assigning Hispanic births. Differences in net international migration assumptions also likely played a role, particularly for older ages.*

Recommendation 4.1: For the Demographic Analysis (DA) coverage evaluation method, the U.S. Census Bureau should conduct research on:

- **The reasonableness of the assumption for incompleteness of Medicare enrollment in the high series, which produced coverage estimates that diverged greatly from the low and middle series for people ages 75 and older, through an appropriate match study;**
- **The suitability of Medicare data for coverage evaluation of the entire population ages 65 and older as in the 2010 Census and for improving the census count of this age group in 2030;**
- **The reasonableness of the methods for assigning Hispanic and non-Hispanic births for coverage estimates for chil-**

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

dren and young adults, which produced wide differences between the low and high DA series in 2020;
- Methods for narrowing estimates of net international migration, which affect estimates for Hispanic people and the total population; and
- Methods for developing experimental subnational DA estimates, starting with young children.

Recommendation 4.2: The U.S. Census Bureau should conduct additional analyses of the 2020 Post-Enumeration Survey (PES) to learn as much as possible about census omissions. One such analysis should include P-sample and E-sample households with at least one member in common, to assess not only erroneous census enumerations but also omissions and how they vary with household characteristics and type of enumeration (e.g., proxy interview, internet response). In addition, the Census Bureau should perform discriminant or similar analyses to identify variables that contributed to census component errors (e.g., omission, duplication) for 2020 P-sample and E-sample addresses, households, and individuals, which could identify census processes to target for improvement in 2030.

Recommendation 4.3: For the Post-Enumeration Survey (PES) method, the U.S. Census Bureau should:

- **Perform sensitivity analyses, based on plausible assumptions, to put error bounds around such key operations as matching and imputation of match and enumeration status, to evaluate the quality of the 2020 PES and plan such analyses from the outset for the 2030 PES;**
- **Plan analyses for the 2030 PES of components of error for census operations in exhaustive and mutually exclusive categories that are as comparable as possible between 2020 and 2030;**
- **Seek adequate funding to increase the PES sample size in 2030 to at least 2010 levels; and**
- **Experiment with modeling to estimate net undercoverage and overcoverage for more detailed geographic and demographic groups for which direct estimates could not be provided in 2020.**

Recommendation 4.4: The U.S. Census Bureau should prioritize research on potential sources of coverage errors—both undercounts and overcounts—for geographic areas and population groups, using additional methods besides further analysis

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

of the Post-Enumeration Survey. To address omissions from the census, the Census Bureau should match 2020 Census records with the 2019–2021 American Community Survey and a wide variety of administrative records for census tracts—perhaps sampling those with low self-response rates. The goal would be to provide an evidence base for methods and data sources that could potentially reduce omissions in the 2030 Census. To address duplications in the census, the Census Bureau should match 2020 Census records with themselves to identify duplicate enumerations of people with more than one residence, college students enumerated at college and at home, children in joint custody arrangements, and the like. The goal would be to provide an evidence base for methods to reduce duplicates in the 2030 Census and potentially to remove them from the count, just as people duplicated within a household are dropped from the count.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

– 5 –

Master Address File

The quality of any sample survey is dependent upon the quality, completeness, and appropriateness of the underlying frame from which the sample is drawn. 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, it follows that the quality of the decennial census as a 100 percent sample of the population is inextricably linked to the quality of the U.S. Census Bureau's geographic resources and, in particular, the address list used to mount census data-collection efforts. Prior to the 2000 Census, the address list used as the operational frame to conduct and monitor progress in the decennial census was constructed anew each decade, prior to and during the census. 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, with MAF entries geocoded (matched to their geographic locations in a census block) using the Topologically Integrated Geographic Encoding and Referencing (TIGER) map database. Since 2000, the MAF has evolved into a fluid, ongoing, regularly updated repository of addresses encountered by the Census Bureau, and the broader MAF is filtered to define operational extracts of MAF units that are potentially "good" or workable for a particular census or survey purpose.

Maintenance and updating of the MAF/TIGER databases has evolved into a high-stakes and critically important enterprise before, during, and after the census. The MAF is updated through a series of suboperations prior to and during each decennial census, including a comprehensive address list check prior to the census, a suite of address-building operations (e.g., Local

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Update of Census Addresses [LUCA] and New Construction [NC] programs[1]) drawing input from census users and stakeholders, and census component operations that can alter MAF entries. Between decennial censuses, the primary update sources of the MAF are twice-yearly "refreshes" of addresses from the U.S. Postal Service's (USPS) Delivery Sequence File (DSF), which is USPS's inventory of residential and nonresidential addresses for mail delivery carrier routes. Information garnered from ongoing fieldwork surrounding Census Bureau surveys (e.g., the American Community Survey) may contribute additional changes to the MAF. An important pillar in the general plan for the 2020 Census was fundamentally retooling the precensus address list review activity, Address Canvassing (AdCan), which had developed into a field-heavy, walk-every-block activity in the 2000 and 2010 Censuses. Instead, a clerical In-Office Address Canvassing (IOAC) component was implemented to bear much of the load of identifying stable, unchanging blocks, in an attempt to limit the traditional In-Field Address Canvassing (IFAC) fieldwork to blocks that could not be resolved by IOAC.

In this chapter, we begin our broader examination of 2020 Census data quality by delving into the mechanics of the 2020 Census MAF. The panel's data analysis subgroup was permitted unprecedented access to a set of operational metadata for every record on the MAF as of the end of the 2020 Census—more than 212 million entries. These data do not include actual addresses; instead, we have Basic Collection Unit (BCU) and census tract identifiers for each record in the dataset. Importantly, the metadata include both the Census Bureau's indication of the address record's provenance—the specific census or update mechanism by which the address was first observed/experienced in the MAF—and an indicator of the record's final status in the 2020 Census (i.e., whether it resulted in an enumeration). We derived a categorization of MAF records to better understand the composition of the 2020 Census MAF, look for patterns by county and tract, and further examine the origin of addresses added or "touched" by the individual address-building operations of the 2020 Census.

Before delving into the composition of the MAF, it is important to recognize that the MAF is intended to be a repository for all living quarters in the United States, including not only housing units but group quarters (GQs) (such as college/university housing, health care/nursing facilities, and correctional facilities) and transitory locations (TLs) (such as marinas, recreational vehicles, and hotels/motels). As we discuss in more detail in Chapter 9, both GQs and TLs have historically been subject to separate frame-building work; in the 2020 Census, both involved "Advance Contact" operations that could impact the ultimate address list for those sites. That said, the available metadata show the

[1]The LUCA Program was created by the Address List Improvement Act of 1994. See: https://www.census.gov/programs-surveys/decennial-census/about/luca/act-1994.html. For information on the NC Program, see: https://www.census.gov/programs-surveys/decennial-census/about/new-construction.html. Also, see: https://www.census.gov/history/pdf/Census2000v2.pdf.

Copyright National Academy of Sciences. All rights reserved.

classification of MAF entries as housing units, GQs, TLs, or nonresidential as of the end of the 2020 Census, but do not provide a good indication of what the addresses were believed to be at prior times. So, while the case could be made for omitting GQs and TLs from the analysis, we do not believe that the data at hand provided a reliable means for doing so. Hence, we retained all address types in our analysis of the MAF (at the expense of some noise in a pure-housing-unit picture), and hence our analyses contain potential discrepancies, specifically in terms of the GQ population.

5.1 OVERVIEW OF MASTER ADDRESS FILE DEVELOPMENT AND THE DATASET

The MAF is a fluid, ongoing, longitudinal database of address information—a record of how addresses for living quarters in the United States (housing units and now including GQ locations) have been represented over time in the Census Bureau's address listings. Addresses are assigned a unique identifier (MAFID) for tracking. Curating an address database is vastly more difficult than it might appear at first glance, particularly given the diversity of sources from which MAF data are drawn. Apartments in small, multifamily dwellings without standard apartment numbers may be represented differently at different times and in different sources (e.g., as Apartment 2, Apartment B, "back," or "basement"); units within large, multiunit structures may be redesignated or renumbered. In rural areas where mail is not delivered to individual homes, standardization and conversion of physical location addresses (as for systems for emergency response and fielding 911 calls) are ongoing processes and can change over time, meaning that different labels might apply to the same physical "house on Lake Bliss" in separate data series. Accordingly, it is sensible that the primary update source for the MAF between censuses is a twice-yearly extract from the USPS DSF for mail carrier routes—not only because of the USPS's daily experience and interaction with mail delivery addresses but also due to its own keen, vested interest in address standardization and general hygiene of address databases.

To conduct the census—or to conduct one of the Census Bureau's other household surveys—a set of filters is applied to the MAF to derive an extract appropriate to the task. In the census context, it is logical that the filters would be made to err on the side of inclusion, to limit the risk of omitting units that should be included and thus maximizing participation in the census. By comparison, a new sampling frame for a particular survey program might purposefully exclude addresses contacted in the past months/years, to avoid burdening survey respondents. However, every address that is included in the working frame for the census necessarily incurs some cost—minimal for mailing but greater when many enumerator/interviewer contact attempts are needed—

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

so including too many questionable addresses on the initial list for census enumeration runs the risk of overwhelming operations and compromising the effectiveness of data collection. Accordingly, to best use its resources, the Census Bureau must try to create a balance between these competing costs and benefits.

To study the MAF development process, the panel's data analysis subgroup was given access to what we refer to as the MAF Development Dataset, consisting of a metadata extract on the 212.3 million addresses (MAF units) that existed in the MAF at the end of 2020 Census data collection (i.e., on October 15, 2020). The dataset contains information on the provenance of the address (the census operation, dating back to the 2000 Census, in which the address was first added to the MAF), the status of the address on the USPS DSF in fall 2017 and spring 2020, and whether action was taken on the address in any of the 15 address-building suboperations of the 2020 Census. For each address, the dataset provides the BCU code which, akin to census blocks, is the smallest unit of collection geography used in the 2020 Census, as well as the census tract codes for 2020 Census tabulation and the older-vintage tracts that prevailed during data collection (as derived from the 2019-vintage BCU code). The dataset also provides important context for our analyses by indicating the type of enumeration area associated with a particular MAF entry's BCU, which designates the 2020 Census' approach to respondent contact and data collection for a particular BCU (see Box 2.2).

For our purposes, the dataset is a rich trove of information because it indicates the outcome of filters run to assess workable addresses at several points: before the beginning of IFAC (run in April 2019), in defining the original collection universe for the 2020 Census (run in November 2019), and in defining the supplemental Nonresponse Followup (NRFU) workload (run in March 2020).[2] The dataset includes the final status[3] of the MAF unit in the 2020 Census—whether it was enumerated as a housing unit by self-response questionnaire or enumerator form, whether it was enumerated as a GQ, whether it was not enumerated (as either a housing unit or GQ), or whether it had no defined status at the end of the census—as well as an associated population count. The final status and population count variables in this dataset are the operational feedback provided to the Census Bureau's Geography Division, and so do not necessarily comport with the final editing,

[2]The supplemental NRFU workload included new addresses added to the NRFU workload that were not in the workload during the initial cut in late 2019/early 2020, such as NC, successfully appealed LUCA addresses, and a late refresh of the USPS DSF. Relevant to the analysis in this chapter, a late change to the NRFU operation in 2020 was an expansion of the supplemental NRFU workload to include units marked "Delete" in Address Canvassing as well as all regular housing units that had been encountered in the initial Advance Visit operations for GQ and TL enumeration (Fontenot, 2021d).

[3]The dataset includes the same final status code for the MAF unit from the 2010 Census.

Copyright National Academy of Sciences. All rights reserved.

imputation, and processing of 2020 Census data. We treat the data items accordingly—not delving too deeply into the status codes beyond "Pass" and "Fail" (enumerated/in the census or not) and distinguishing between occupied, zero-population, and null/missing counts.

5.2 MAF1 AND MAF2, PRECENSUS AND POSTCENSUS PIVOT POINTS IN THE MASTER ADDRESS FILE

To understand the dynamics in the development of the MAF, our starting point was to echo and extend the work of our panel's predecessors on these lines, who performed data analysis (Biemer et al., 2021) for the American Statistical Association's (ASA) 2020 Census Quality Indicators Task Force (2020, 2021). We focused on two key pivot points in the development of the MAF for the 2020 Census. In this framework, MAF1 is the MAF extract constructed at the beginning of 2020 Census data-collection operations (i.e., as of March 2020) and MAF2 is the operational address frame at the end of 2020 Census data collection (i.e., as of October 15, 2020). Figure 5.1 includes both a simple Venn diagram of these two pivot points and a schematic of how they are operationalized in the MAF Development Dataset: MAF1 as the set of addresses passing the filter and deemed workable addresses in either the initial 2020 Census frame created in November 2019 or the supplemental workload created in March 2020;[4] and MAF2 derived from the variable (CENSTAT2020) on the dataset indicating whether the MAF unit yielded a housing unit or GQ enumeration in the 2020 Census.

The general MAF1 and MAF2 framework is sound, with before-2020-collection and after-2020-collection being critically important pivot points for the MAF. However, while the ASA team considered a simple "MAF Revisions" statistic as a measure of churn in the MAF—the proportion of all addresses in the combined MAF1 and MAF2 sets that are not listed on both MAF1 and MAF2, or the sum of MAF1 Only and MAF2 Only in the Figure 5.1 Venn diagram—we concluded that the approach required additional finesse. The set of addresses that are in the MAF1 Only set arguably represent bad outcomes: addresses deemed workable just before the census but that did not yield a successful enumeration, thus representing an investigative dead-end and some squandering of resources. But the set of addresses that are MAF2 Only—successfully enumerated in the census, but not in the initial census workload—are not only good outcomes but very much expected and desired

[4]The two filter variables, coded as ENUMUNV and SUPNRFUUNV in the dataset, are properly thought of as a two-part filter, with the supplemental universe purposefully emphasizing inputs from new-since-late-2019 address sources such as LUCA and the most recent DSF refresh in spring 2020. That is, the two variables are not meant to replicate each other—an address passing the first cut but not passing the supplemental cut is not inherently problematic—and so MAF1 is deemed a "Pass" if the MAF unit passes either ENUMUNV or SUPNRFUUNV.

Copyright National Academy of Sciences. All rights reserved.

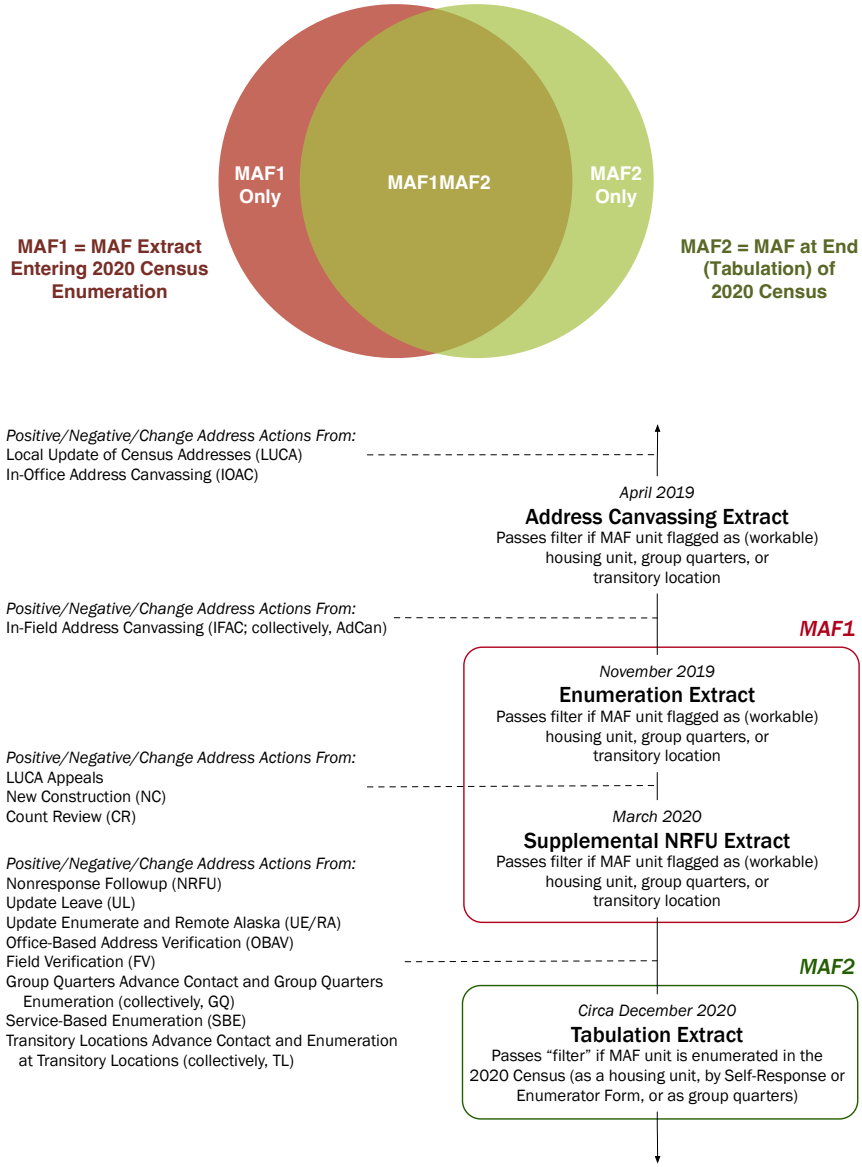


Figure 5.1 Basic formulation of MAF1 (entering 2020 Census enumeration) and MAF2 (2020 Census tabulation) filters.

SOURCE: Panel generated.

Copyright National Academy of Sciences. All rights reserved.

Table 5.1 MAF1 (Entered Enumeration Process in 2020 Census) and MAF2 (Actual Enumeration in 2020 Census) Filters

	MAF2 (Enumerated in 2020 Census)			
MAF1 (Enter Enumeration)	Pass	Fail	Unknown	Total
Pass	140,100,000	11,460,000	747,000	152,300,000
Fail	742,000	72,500	57,300,000	58,110,000
Unknown	1,474,000	328,000	84,000	1,886,000
Total	142,300,000	11,860,000	58,130,000	212,300,000

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0215.

outcomes. In the non-Self-Response Update Leave and Update Enumerate areas of the nation, and in other areas where user/stakeholder input through address update operations contributed good, new addresses, the discovery of heretofore unknown-to-the-MAF addresses during the census itself is beneficial. Hence, simply adding the MAF1 Only and MAF2 Only slivers to form a single statistic is misleading.

In arriving at a partition of addresses, there is an important subtlety: each of the MAF1 and MAF2 pivot points has three basic outcomes, not just two as represented in the Venn diagram. In the case of MAF1, an address can pass the filter and be deemed workable (Pass), explicitly fail the filter and be excluded from initial workloads (Fail), or have null status because the address does not exist or is not yet on the MAF (Unknown). Likewise, the MAF2 filter that corresponds to final status in the 2020 Census can be simplified to a "pass" or enumeration in the census (Pass) or a "fail" indicating failure to be enumerated in the census (Fail), but may also have no status or unknown/null status at the end of the census (Unknown). The full set of addresses in the dataset is distributed across the nine possible cells, as shown in Table 5.1.

The largest single bloc of addresses in the table is the MAF1–MAF2 intersection event (140.1 million), and the MAF1 Only and MAF2 Only slivers are the remnants of the MAF1 "Pass" row and MAF2 "Pass" column, respectively. It is more difficult to attach meaning to the collection of addresses that are designated "Fail" or "Unknown" in both the MAF 1 and MAF2 filters. There is a temptation to consider these addresses as "dead weight" on the MAF, but that may be an overly hasty conclusion; those addresses are arguably benign if the filtering process works and there is no real investment of time or resources in trying to enumerate them (i.e., if they are truly held out of the census collection process). Accordingly, we turn to another variable coded in the dataset to split the (MAF1 and MAF2 both in {Fail, Unknown} into two pieces:

Copyright National Academy of Sciences. All rights reserved.

- Roughly 431,000 addresses for which COLLUNV = *Y* meaning that some 2020 Census collection effort was made, a group that we label "Remnant in Collection;" and
- About 57,350,000 MAF entries for which COLLUNV = *N*, which we label as "Extraneous" to the 2020 Census collection universe.[5]

This Extraneous category contains most, but not all, of roughly 12.2 million addresses flagged in the MAF Development Dataset as known duplicate entries; it should also be noted that there are almost certainly more duplicate entries in the MAF than the 12.2 million already identified by the Census Bureau, but we have no way of ascertaining those.

The Extraneous category may require some reflection because it represents potential "dead weight" or drag on the MAF, and arguments can be made for either retaining or culling a large chunk of these addresses. However, in the remainder of this report, we set the Extraneous addresses aside and work with the remaining 154.9 million addresses in the effective 2020 Census data-collection universe.

5.3 PROVENANCE OF ADDRESSES IN THE 2020 CENSUS MASTER ADDRESS FILE AND RELATION TO ADDRESS-BUILDING OPERATIONS

The SRCCAT variable in the MAF Development Dataset is the Census Bureau's assessment of the provenance of each MAF entry—the single operation through which the address was first added to the MAF. These data include codes for 26 provenance categories, which we have condensed to 3 broad categories and 16 subcategories:

- *Legacy Censuses*, including subcategories for the 2000 Census MAF, the 2010 Census MAF, and census research activities (2018 End-to-End Test, MAF Coverage Study, and other recurring or single-occurrence operations) predating the 2020 Census;
- *Legacy DSF*, leaving the uniquely important immediate-precensus spring 2020 refresh from the DSF as its own subcategory but consolidating source categories from seven earlier DSF refreshes into a single group; and
- *2020 Census Address-Building Operations*, with subcategories for each of 11 suboperations of the 2020 Census that had the capability for adding or revising addresses.

[5]This calculation involves correcting for some inconsistencies—notably, a reasonably small group of addresses simultaneously coded as COLLUNV = *N* (no enumeration attempted) and as a MAF2 pass (successful enumeration in the census), in which we take the latter condition to override the former. As with other analyses we have performed, sums may differ between and within tables due to the rounding and disclosure avoidance required by the Census Bureau.

Copyright National Academy of Sciences. All rights reserved.

Under the 2020 Census Address-Building Operations heading, "adds" to the MAF can come from 11 distinct operations:

- Address Canvassing (AdCan), combining separate origin codes for IOAC and IFAC;[6]
- The three principal operations through which user/stakeholder input is brought to bear on MAF content—LUCA, Count Review (CR), and NC;
- Office-Based Address Validation (OBAV), new to the 2020 Census, which extended IOAC methodologies to try to resolve addresses from Non-ID Processing that did not automatically match to the MAF;
- Field Verification (FV), or field work on addresses from Self-Response and other operations that could not be resolved in-office (in the 2020 Census, FV additions to the MAF were exclusively made in Puerto Rico);
- NRFU, the core field follow-up operation of the census, in which enumerators could add an address observed during their work;
- The two non-Self-Response enumeration strategies, Update Leave (UL) and Update Enumerate/Remote Alaska (UE/RA),[7] in which the "Update" part is particularly important because these areas were excluded from IFAC ahead of the census; and
- Special enumeration operations for GQ and TL, both of which included advance-visit suboperations that could determine new addresses as well as broker arrangements for the eventual enumeration.

Cross-tabulating our working MAF composition variable with the SRCCAT address provenance variable yields Table 5.2, the most striking observation from which is arguably the sheer magnitude of addresses originating in the 2000 Census and in pre-2020 iterations of the DSF. The 2000 Census MAF alone accounts for the origin of 57.2% (88.6 million) of the addresses in the 2020 Census data-collection universe; the pre-2020 DSF adds another 32.6% (50.5 million) and the 2010 Census contributes 5.1% (7.9 million). The result is understandable given the semipermanent nature of housing stock, yet it can still be surprising to realize that less than 5% of the addresses in the 2020 Census data-collection universe were uniquely added by 2020 Census operations. This is decidedly not meant to suggest that those 2020 Census address-building operations are unimportant—they definitely are important to the quality of the count—but just meant as calibration, to better understand continuity in the MAF.

Comparing the first two columns of Table 5.2, MAF1MAF2 and MAF1 Only, suggests a fundamental difference between the various address provenance

[6]These are almost exclusively addresses impacted by IFAC, because the Active Block Resolution component of IOAC that would have permitted additions/revisions in the in-office phase was suspended early for budgetary reasons.

[7]Update Enumerate and Remote Alaska are coded as separate variables in the dataset, but they are combined because of their similar methodology.

Copyright National Academy of Sciences. All rights reserved.

Table 5.2 Master Address File Entries by Typology Based on MAF1/MAF2/Collection Attempt Flags and Address Origin

Address Origin	MAF1MAF2	MAF1 Only	MAF2 Only	Remnant in Collection	Total
Legacy Censuses					
2000 Census	84,580,000	3,816,000	186,000	21,000	88,600,000
2010 Census	5,715,000	2,067,000	63,500	10,000	7,856,000
Other Pre2020	697,000	239,000	7,100	1,000	944,000
Legacy DSF					
Pre2020	46,850,000	3,139,000	462,000	45,000	50,500,000
Spring 2020	146,000	74,500	22,000	3,900	246,000
2020 Census Operations					
LUCA	1,216,000	2,020,000	1,400	250	3,238,000
CR	40,500	67,500	100	20	108,000
NC	192,000	306,000	60	0	497,000
AdCan	633,000	482,000	80	20	1,115,000
OBAV	0	0	721,000	75,000	795,000
FV	0	0	6,400	0	6,400
NRFU	$N < 15$	50	174,000	25,000	199,000
UL	0	0	401,000	136,000	538,000
UE/RA	0	0	7,300	40	7,400
GQ	0	0	64,500	37,500	102,000
TL	0	0	98,500	76,000	174,000
Total	140,100,000	12,210,000	2,216,000	431,000	154,900,000

NOTES: DSF, Delivery Sequence File; LUCA, Local Update of Census Addresses; CR, Count Review; NC, New Construction; AdCan, Address Canvassing; OBAV, Office-Based Address Verification; FV, Field Verification; NRFU, Nonresponse Followup; UL, Update Leave; UE/RA, Update Enumerate/Remote Alaska; GQ, Group Quarters; TL, Transitory Locations.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0215.

categories. Recall that MAF1 means passing the filter to be deemed workable at the start of the 2020 Census, so MAF1MAF2 connotes those "good" precensus addresses that resulted in successful enumeration, and MAF1 Only denotes those that did not result in enumeration—those addresses that did not pan out and might be considered fruitless investigative leads. By large margins, the share of older-source addresses in MAF1MAF2 (deemed workable and successfully enumerated) dwarfs that in MAF1 Only (deemed workable but not enumerated). This suggests that the filters worked reasonably well for these older-address categories, with addresses identified as workable generally resulting in enumerations. However, of the 2020 Census suboperations, only Address Canvassing worked in that same way with good/enumerated addresses outweighing fruitless leads—and then only by a slight margin. Addresses provided by the three user-input operations (LUCA, CR, and NC) showed

Copyright National Academy of Sciences. All rights reserved.

different enumeration outcomes compared with the other sources. Nearly twice as many new addresses added by LUCA failed enumeration, and majorities of the CR and NC additions likewise failed. Thus, it appears as if the Census Bureau opted to place new, locally provided addresses into the census operational workload and let fieldwork sort out the results. This is neither meant as undue praise of the MAF1 filter nor as undue criticism of the user-input programs—9.3 million fruitless older-source leads and 1.5 million good adds from user-input 2020 Census operations are both sizable blocs, after all—but just an illustration of the trade-offs at play.

The entries in the MAF2 Only column mainly behave as expected. As addresses that were not in the MAF1 filter (initial 2020 Census universe) but did result in successful enumerations, the MAF2 Only category is mainly comprised of addresses coming in through the 2020 Census address-building operations. OBAV was a particularly strong contributor of new-to-MAF and successfully enumerated addresses, at about 721,000; large shares also came in through UL and NRFU. But, again, it is the long-lived nature of the MAF that draws the eye in this category: about one-third of the addresses in the MAF2 Only category originated in previous censuses or DSFs. Logically, this means that the addresses had to: (1) have been added by an older operation; (2) become dormant at some point in terms of being deemed workable and failed the initial filter; and (3) become "reactivated" somehow—making it into the 2020 Census by some means—and resulting in a successful enumeration. Below, we will consider whether we can recover any information about how this reactivation came about.

To sharpen the contrasts somewhat, we developed a final typology in Table 5.3, combining elements of Table 5.2 differently to try to form more meaningful groups. The rows of Table 5.3 mirror those in Table 5.2, but the columns differ:

- *Bad 2020 Adds*: These entries were added to the MAF by a particular 2020 Census address-building operation but were not successfully enumerated.
- *Bad Legacy*: These entries were added to the MAF by a legacy operation, including the 2000 and 2010 Censuses or various DSFs, but were not successfully enumerated.
- *Good 2020 Adds*: These entries were added to the MAF by a particular 2020 Census address-building operation and were successfully enumerated.
- *Good Legacy Census*: These entries were added to the MAF during the 2000 or 2010 Censuses or pre-2020 Census work and were successfully enumerated.
- *Good Legacy DSF*: These entries were added to the MAF from various versions of the DSF and were successfully enumerated.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

- *Good Reactivations*: These entries were added to the MAF by a legacy operation, including the 2000 and 2010 Censuses or various DSFs. They did not pass the original enumeration universe filter(s), but they were reactivated by a 2020 Census address-building operation (e.g., LUCA, AdCan, NRFU, etc.) and resulted in a successful enumeration.

The Legacy Census and DSF sources contributed 138 million MAF entries that were successfully enumerated, with 2020 Census address-building operations contributing 3.5 million successful enumerations. If we compare the outcomes for MAF entries added by 2020 address-building operations, just over half resulted in an enumeration. While a large fraction of adds from 2020 address-building operations was not enumerated, the number of Legacy MAF entries that were not enumerated was three times as large (9.4 million vs. 3.2 million). While most legacy entries resulted in successful enumerations (e.g., 84.6 of 88.6 million addresses from the 2000 Census), the number of legacy entries not enumerated is still large and provides support for continued research into optimizing the filters used to create the initial enumeration and the supplemental NRFU universes.

Examination of the good and bad outcomes for 2020 Census address-building operation additions may lead to the conclusion that such operations are inefficient because just over 50% of the records yield no enumeration. There are two important observations to make about addresses that were the result of census operations. As was evident in Table 5.2, it bears repeating that the balance of successful to unsuccessful enumerations was very different for addresses that came in via user programs versus those that came in via census address-building operations, with the latter having a much better success rate. Second, the values in the Good Reactivations column provide further evidence of the utility of the 2020 Census address-building operations. MAF entries in the Good Reactivation category were not originally included in the 2020 enumeration universe (i.e., they were not in MAF1), but they were "found" by a 2020 Census address-building operation and successfully enumerated. We discuss the operational "touches" of 2020 address-building operations in more detail in Section 5.5, but it is useful to note that over 80% of the 742,000 MAF units in the Good Reactivations category were touched by only one 2020 Census address-building operation; hence, it is easy to surmise that this single operation is the one that "found" the address. Of those single-touch reactivation cases, over 60% were touched solely by the OBAV operation and over 20% were touched solely by NRFU.[8] Without these operations—OBAV verifying Non-ID addresses and NRFU doing limited investigation of addresses at the same time as garnering responses, even amidst a compressed timeframe—the census would have missed those units, which again argues for increased attention to

[8]Of those Good Reactivation cases touched by only two 2020 operations, the majority were touched by OBAV in combination with either Address Canvassing, NRFU, or LUCA.

Copyright National Academy of Sciences. All rights reserved.

Table 5.3 Disposition of 2020 Census Collection Universe Master Address File Entries by Address Origin and Final 2020 Census Status

Address Origin	Bad 2020 Adds	Bad Legacy	Good 2020 Adds	Good Legacy Census	Good Legacy DSF	Good Reactivations	Total
Legacy Censuses							
2000 Census	—	3,837,000	—	84,580,000	—	186,000	88,600,000
2010 Census	—	2,077,000	—	5,715,000	—	63,500	7,856,000
Other Pre2020	—	240,000	—	697,000	—	7,100	944,000
Legacy DSF							
Pre2020	—	3,185,000	—	—	46,850,000	462,000	50,500,000
Spring 2020	—	78,500	—	—	146,000	22,000	246,000
2020 Census Operations							
LUCA	2,020,000	—	1,218,000	—	—	—	3,238,000
CR	67,500	—	40,500	—	—	—	108,000
NC	306,000	—	192,000	—	—	—	497,000
AdCan	482,000	—	633,000	—	—	—	1,115,000
OBAV	75,000	—	721,000	—	—	—	795,000
FV	0	—	6,400	—	—	—	6,400
NRFU	25,000	—	174,000	—	—	—	199,000
UL	136,000	—	401,000	—	—	—	538,000
UE/RA	40	—	7,300	—	—	—	7,400
GQ	37,500	—	64,500	—	—	—	102,000
TL	76,000	—	98,500	—	—	—	174,000
Total	3,225,000	9,416,000	3,556,000	90,990,000	47,000,000	742,000	154,900,000

NOTES: DSF, Delivery Sequence File; LUCA, Local Update of Census Addresses; CR, Count Review; NC, New Construction; AdCan, Address Canvassing; OBAV, Office-Based Address Verification; FV, Field Verification; NRFU, Nonresponse Followup; UL, Update Leave; UE/RA, Update Enumerate/Remote Alaska; GQ, Group Quarters; TL, Transitory Locations.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0215.

Copyright National Academy of Sciences. All rights reserved.

and further refinement of addresses at the MAF1 stage of the census address-selection process.

5.4 SPATIAL VARIATION IN MASTER ADDRESS FILE UNIT CATEGORIES

Table 5.3 provides a national overview of the contribution of various sources to successful and unsuccessful enumerations, but it necessarily obscures spatial variation in the various categories. While some categories of addresses in Table 5.3 may seem small relative to the overall number of addresses in the census, they may nonetheless be important because of differential impacts due to spatial clustering. To examine this spatial variation, we created a series of maps by county and census tract.

5.4.1 Legacy Sources

Figure 5.2 shows the percent of MAF units in the Good Legacy Census category (MAF units from the 2010 and 2000 Censuses or pre-2020 Census operations) by county. In many U.S. counties, more than 50% of their MAF units derive from legacy censuses. Rust Belt and Great Plains counties frequently derive more than 70% of their MAF units from legacy censuses. In the Southeast, it is much more common for counties to have less than 50% of MAF units originating from legacy censuses.

The census tract level provides more fine-grained spatial variation (Figure 5.3). Legacy censuses account for large fractions of MAF units in rural and central city parts of Iowa. In the suburbs around Iowa's larger cities, legacy censuses were the source of a smaller fraction of MAF units. Suburban areas may have been developed after the 2010 Census; therefore, they would not have MAF units from legacy censuses.

Figure 5.4 shows the percent of MAF units in the Good Legacy DSF category by county. This map somewhat mirrors the Good Legacy Census map (Figure 5.2). In many U.S. counties, the DSF provided less than 50% of a county's MAF units. The DSF was an important source of MAF units in the Southeast. Drilling down to the census tract level in the Austin, Texas metropolitan area (Figure 5.5), we observe many census tracts with more than 75% of MAF units originating from the DSF.

Taken together, these four maps highlight the contributions of legacy censuses and the DSF to the MAF. Legacy censuses provide a bulk of the MAF entries that were successfully enumerated in the 2020 Census, and the DSF is a key source of MAF entries in particular areas of the United States. However, the maps show variation by area of the nation, with the south and especially the southeast having smaller shares of their MAF units from legacy sources,

Copyright National Academy of Sciences. All rights reserved.

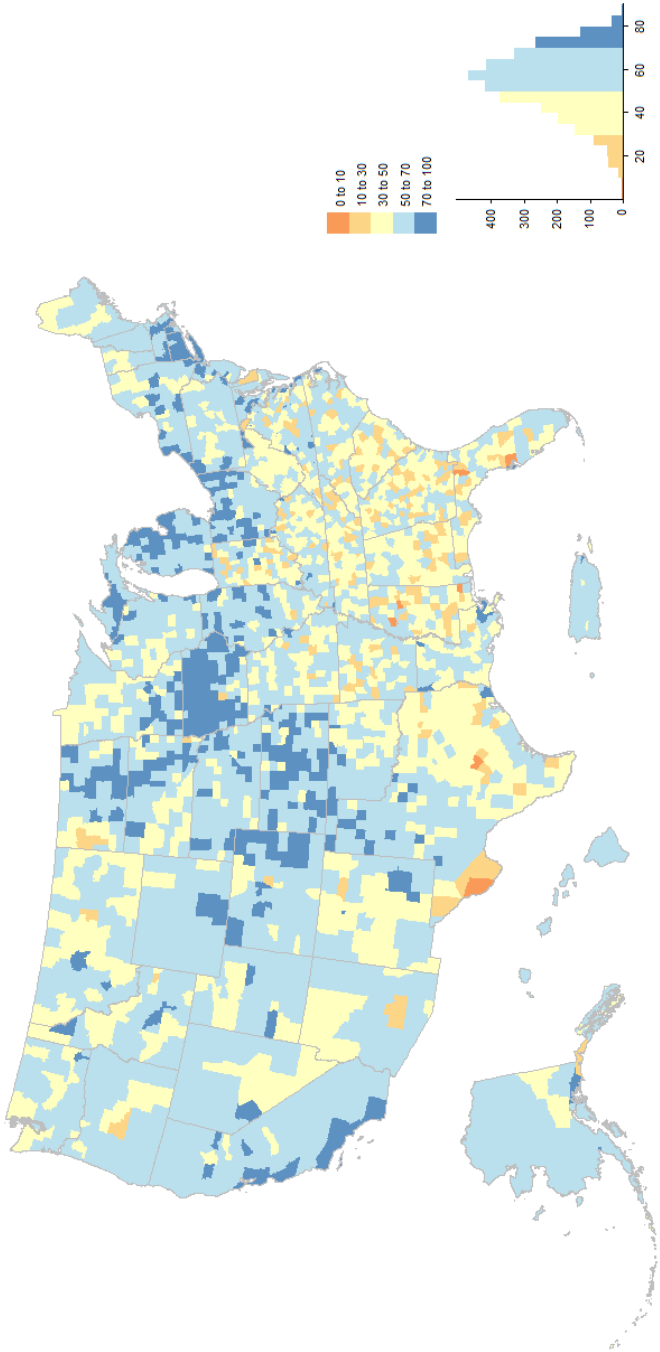


Figure 5.2 Percent of Master Address File units in Good Legacy Census category, by county.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

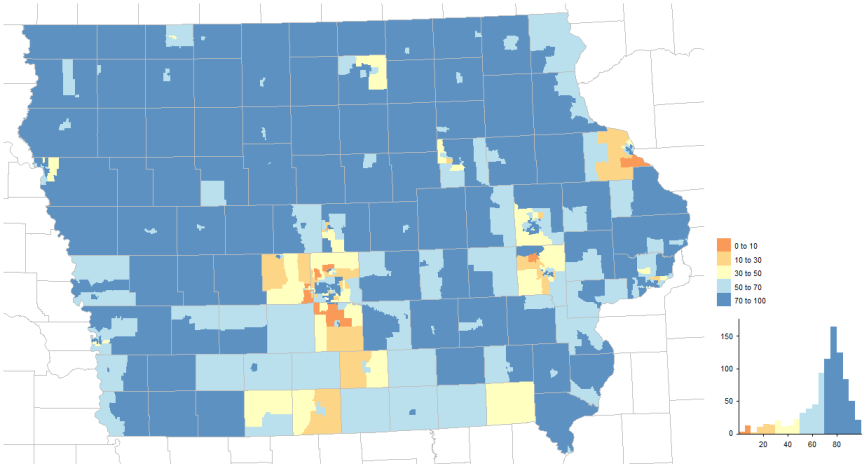


Figure 5.3 Percent of Master Address File units in Good Legacy Census category, by tract, Iowa.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

which translates into a higher share of MAF units from census address-building operations.

It is interesting to see the contrast between the Good Legacy maps and the map of concentration of Bad Legacy addresses shown in Figure 5.6. Pockets of high concentration of Bad Legacy addresses are relatively few and widely scattered, and examination of finer-grained maps (not printed here) point to several reasons why segments of "old," long-known addresses could suddenly fail to show up as 2020 Census enumerations—for instance, areas impacted by natural disasters or likely conversions (or changes in designation) between housing units and GQ locations. At the national/county level, the banding of low concentrations of these Bad Legacy cases in the Rust Belt, Mid-Atlantic, and California and Florida coasts is striking.

5.4.2 Good Reactivations

When we consider the spatial distribution of MAF units that were "reactivated" by a 2020 Census address-building operation and then successfully enumerated, we observe a few counties with more than 4% of their MAF units reactivated (Figure 5.7). This is not a huge surprise given the relatively small number of records in this category. At the local scale, however, we do observe multiple census tracts in the Dallas-Fort Worth metropolitan area with greater

Copyright National Academy of Sciences. All rights reserved.

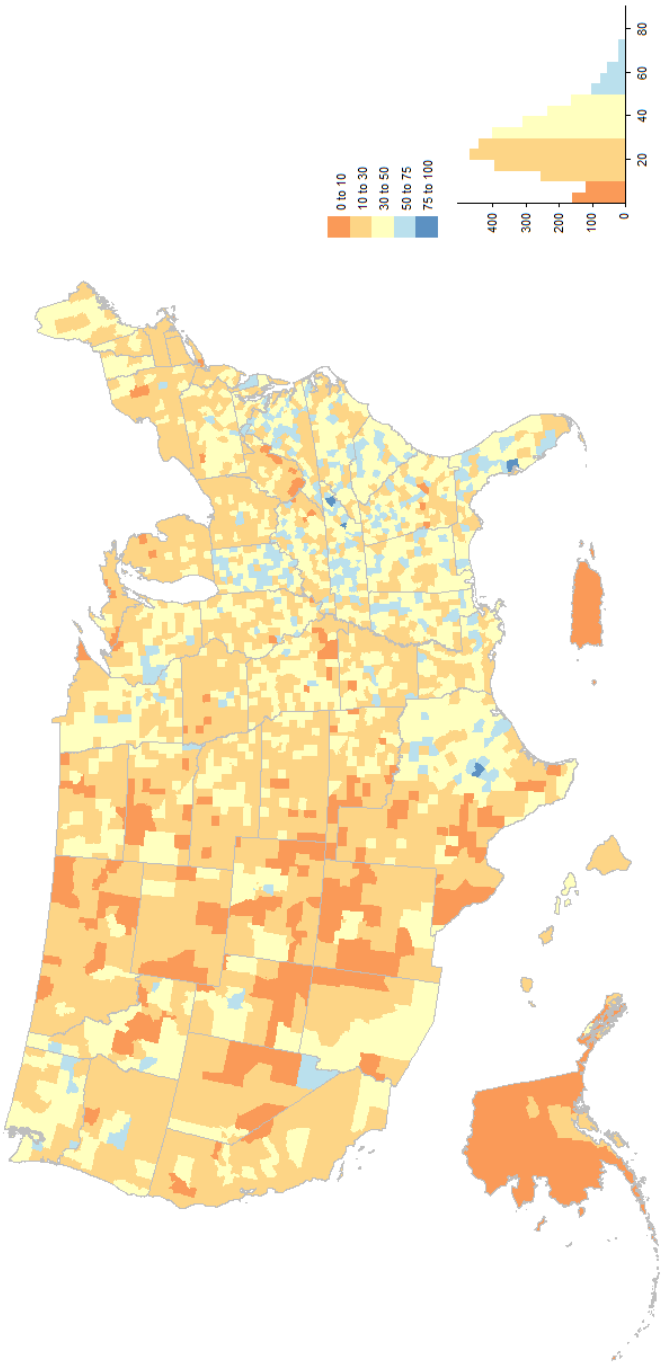


Figure 5.4 Percent of Master Address File units in Good Legacy Delivery Sequence File category, by county.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

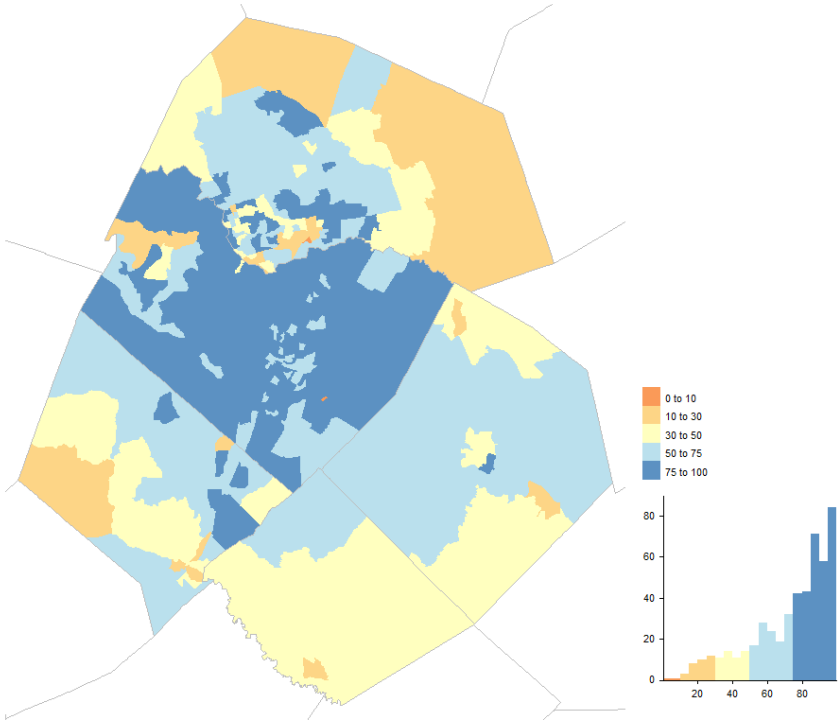


Figure 5.5 Percent of Master Address File units in Good Legacy Delivery Sequence File category, by tract, Austin, Texas metropolitan area.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

than 20% of their MAF units classified as reactivated (Figure 5.8). These are areas that may have been undercounted had census address-building operations not successfully identified and enumerated MAF units excluded from the initial enumeration universe.

5.4.3 Unique Adds from 2020 Census Address-Building Operations

Figure 5.9 depicts county-level concentrations of "good" (enumerated in 2020) MAF units uniquely sourced on the MAF to a 2020 Census address-building operation, and Figure 5.11 does the same for "bad" (not enumerated in 2020) 2020 adds; Figures 5.10 and 5.12 show tract-level concentrations for the same section of southern California, spanning Los Angeles, Orange, and

Copyright National Academy of Sciences. All rights reserved.

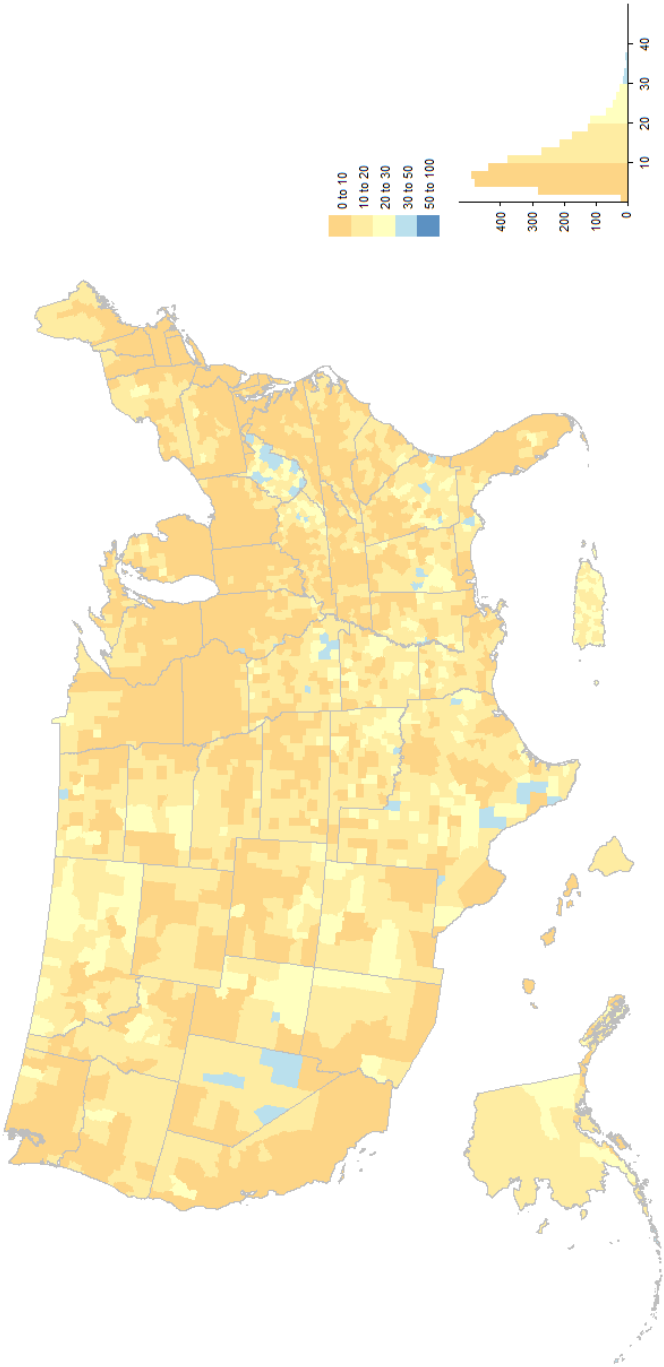


Figure 5.6 Percent of Master Address File units in Bad Legacy category, by county.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

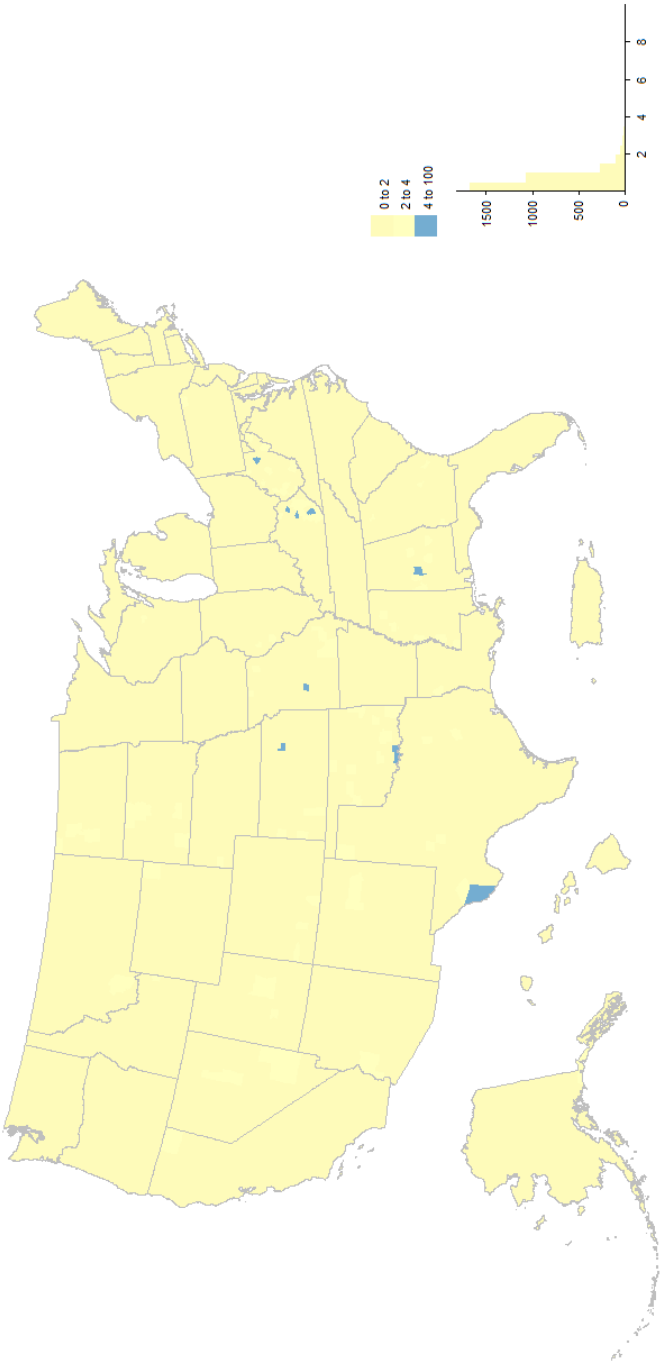


Figure 5.7 Percent of Master Address File units in Good Reactivations category, by county.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

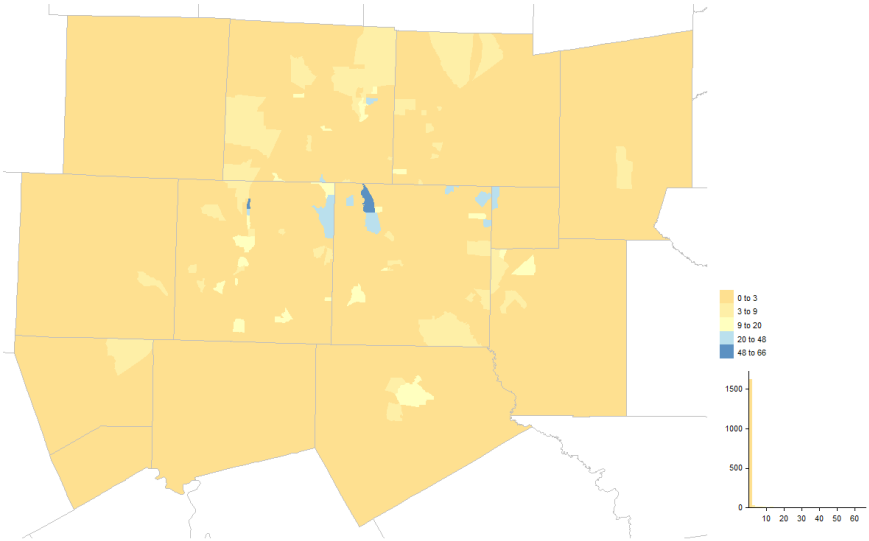


Figure 5.8 Percent of Master Address File units in Good Reactivations category, by tract, Dallas–Fort Worth metropolitan area.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

San Diego Counties. The interesting picture is how few and localized the high concentrations are. The 2020 Census address-building operations vary in terms of participation and coverage, and the results defy easy summarization. For instance, it is an eclectic mix of tracts comprising current and former military facilities, colleges and universities, parks, and airports that show up as high concentrations in *both* the "good" and "bad" 2020 adds maps for southern California, but with unique subsets between the two maps.

5.5 IMPACT OF 2020 CENSUS ADDRESS-BUILDING OPERATIONS

The prior discussion focused on the origin of MAF entries and their final disposition during the 2020 Census. We highlighted the fact that contributions from user-input address-building operations (LUCA, NC, and CR) and other 2020 Census address-building operations (e.g., AdCan, NRFU, UL) compose a small portion of the final 2020 enumeration. Additionally, with few exceptions (e.g., AdCan, OBAV, NRFU, UL), MAF entries contributed by these operations were more likely than not to fail enumeration.

Copyright National Academy of Sciences. All rights reserved.

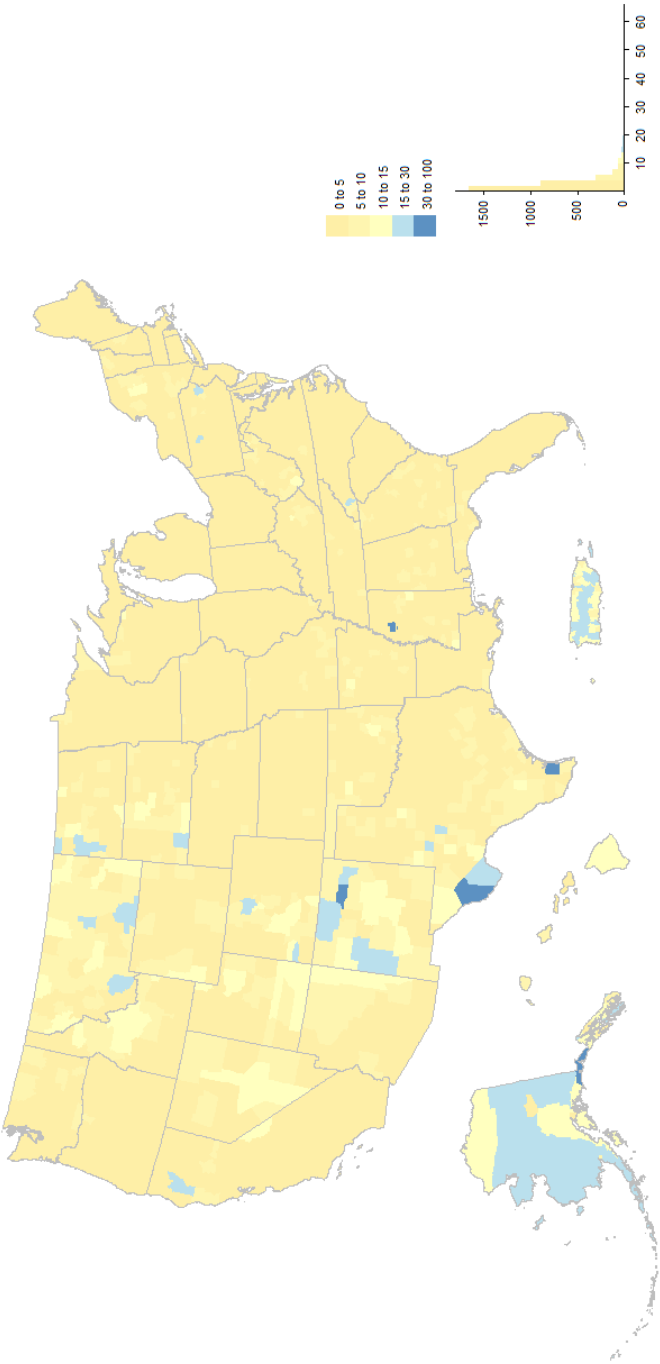


Figure 5.9 Percent of Master Address File units in Good 2020 Adds category, by county.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

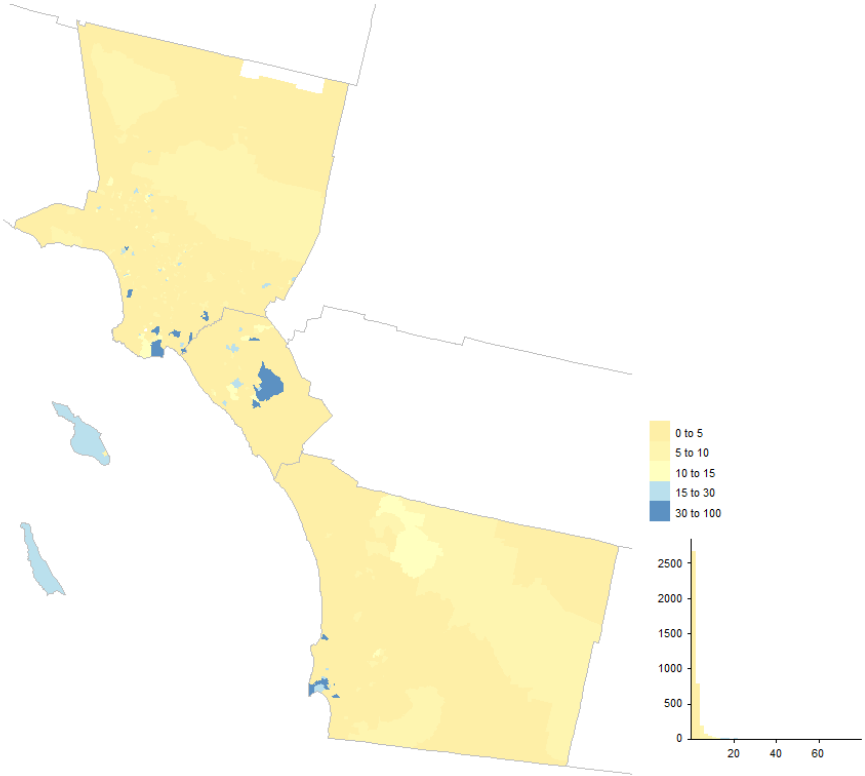


Figure 5.10 Percent of Master Address File units in Good 2020 Adds category, by tract, Los Angeles–Orange–San Diego Counties, California.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

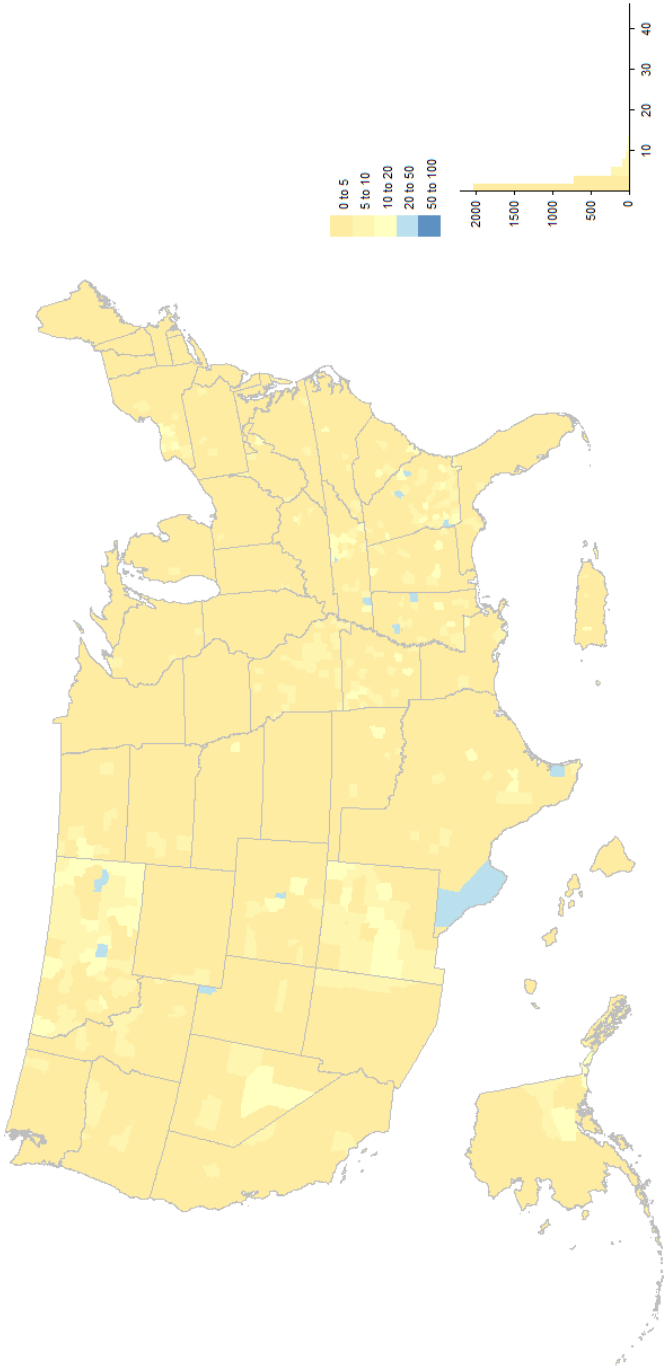


Figure 5.11 Percent of Master Address File units in Bad 2020 Adds category, by county.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

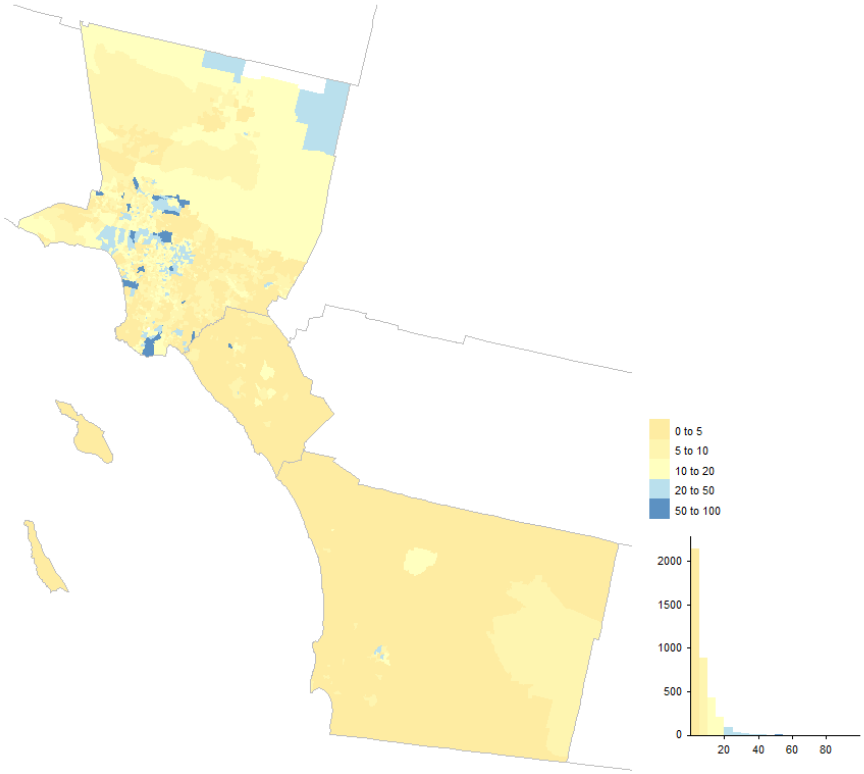


Figure 5.12 Percent of Master Address File units in Bad 2020 Adds category, by tract, Los Angeles–Orange–San Diego Counties, California.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

These operations, though, interacted with the MAF more deeply than is indicated by Tables 5.2 and 5.3. The MAF Development Dataset provides indicators for whether particular 2020 Census address-building operations took an action on a MAF entry, as well as the origin of the address. Thus, we can determine whether a particular address-building operation (1) added an address to the MAF entry or (2) took an action (or "touched") a MAF entry.[9] Tabulating adds or actions by address origin provides us with Table 5.4, which is segmented by major category of 2020 Census address-building operations:

- User Input Operations—LUCA, NC and CR;
- Field/Clerical Operations—AdCan, OBAV, FV, and NRFU; and
- Enumeration Operations—UL, UE/RA, GQ, TL, and Service-Based Enumeration.[10]

All told, 2020 Census suboperations added or touched 83.9 million MAF entries.

5.5.1 Address List Input from User Programs

While LUCA, NC, and CR had mixed results in terms of successful enumerations, some areas of the nation had large fractions of their addresses affected or "touched" in some way by these user-input programs. LUCA was the most influential of these programs. There were 16.5 million MAF units submitted by LUCA participants, 12 million of which had origins in previous censuses and pre-2020 DSF updates, meaning that just 3.2 million originated uniquely from the 2020 Census LUCA program.

In parts of Montana, Nevada, Colorado, New Mexico, and Texas as well as clusters in the Midwest, Northeast, and South, 30% or more of all MAF units were touched in some way by user-input programs (Figure 5.13). Perhaps the greatest concentration in the nation can be found in the District of Columbia where the large majority of addresses were touched by LUCA actions (Figure 5.14). Other examples of marked local concentrations of addresses affected by LUCA can be found in Florida, in the Tampa-St. Petersburg and Miami-Dade County metropolitan areas (Figures 5.15 and 5.16).

[9]The flags in the dataset indicate whether the action taken was "positive," "negative," or "change," but we chose not to distinguish between the types. The principal reason for doing so was to avoid the near-automatic association between "positive" action and "add" and between "negative" action and "delete," when neither of those associations is exactly apt. For example, the neutral action of simple verification of the existence of an address could be counted as a positive action in the data; likewise, a negative action in one address-building operation does not necessarily mean automatic deletion.

[10]Service-Based Enumeration is the sole 2020 Census address-building operation for which the MAF Development Dataset provides an action flag but that does not appear as a separate entry in the address origin (SRCCAT) variable.

Copyright National Academy of Sciences. All rights reserved.

Table 5.4 Addresses Added or Touched by 2020 Census Address-Building Operations, by Address Origin

	2020 Census User-Input Operation		
Address Origin	LUCA	CR	NC
Legacy Censuses			
2000 Census	5,501,000	23,500	173,000
2010 Census	865,000	14,500	37,500
Other Pre2020	83,000	300	4,600
Legacy DSF			
Pre2020	6,856,000	93,500	779,000
Spring 2020	0	0	0
2020 Census Operation			
LUCA	3,238,000	100	65,500
CR	0	108,000	4,300
NC	0	0	497,000
AdCan	400	0	36,500
OBAV	0	0	0
FV	0	0	0
NRFU	0	0	0
UL	0	0	0
UE/RA	0	0	0
GQ	0	0	0
TL	0	0	0
Total	16,540,000	240,000	1,599,000

	2020 Census Field/Clerical Operation			
Address Origin	AdCan	OBAV	FV	NRFU
Legacy Censuses				
2000 Census	21,880,000	295,000	12,500	1,778,000
2010 Census	2,834,000	66,500	2,100	838,000
Other Pre2020	286,000	6,700	300	76,000
Legacy DSF				
Pre2020	22,320,000	390,000	10,500	1,852,000
Spring 2020	0	13,000	350	64,500
2020 Census Operation				
LUCA	1,325,000	4,000	400	687,000
CR	0	350	$N < 15$	35,000
NC	0	2,500	50	153,000
AdCan	1,115,000	1,000	80	239,000
OBAV	0	795,000	94,000	3,000
FV	0	6,400	6,400	$N < 15$
NRFU	0	1,200	200	199,000
UL	0	8,600	4,000	53,500
UE/RA	0	20	$N < 15$	$N < 15$
GQ	0	$N < 15$	$N < 15$	0
TL	0	$N < 15$	$N < 15$	$N < 15$
Total	49,760,000	1,590,000	131,000	5,977,000

Copyright National Academy of Sciences. All rights reserved.

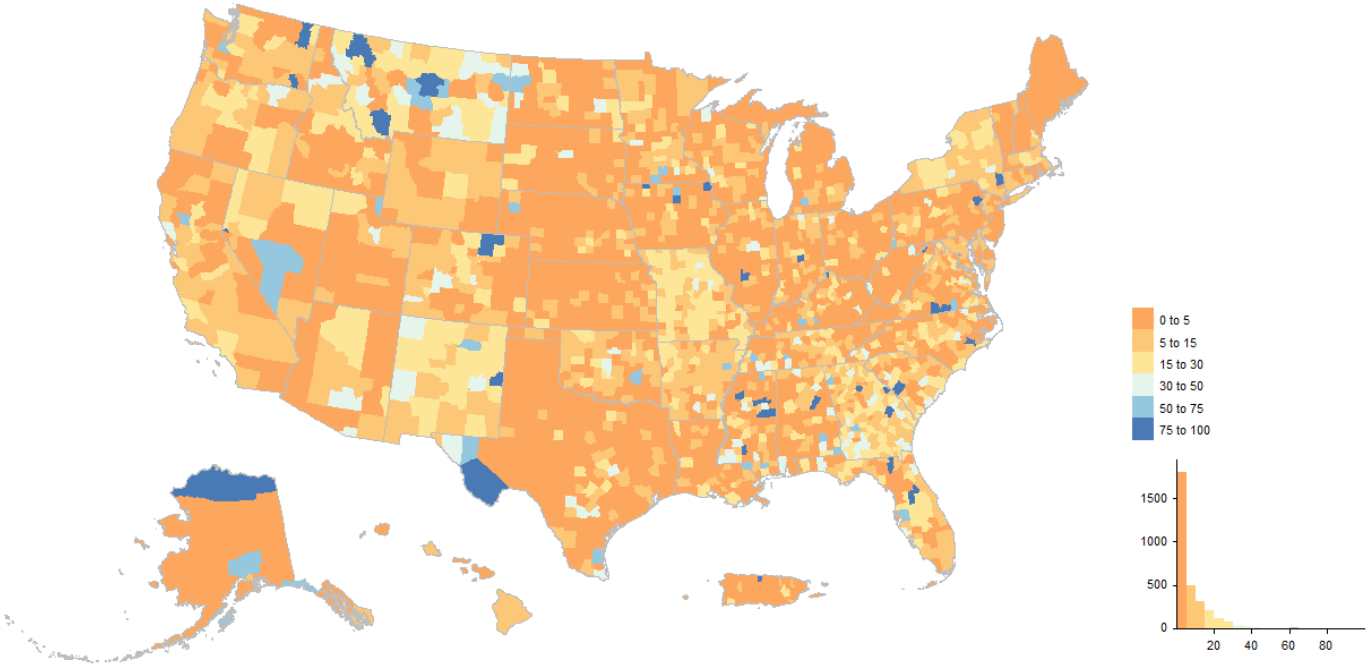


Figure 5.13 Percent of Master Address File units added or touched by 2020 Census Local Update of Census Addresses operation, by county.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

Table 5.4 (cont.)

	2020 Census Enumeration Operation					
Address Origin	UL	UE/RA	GQ	TL	SBE	Total (all operations)
Legacy Censuses						
2000 Census	3,097,000	21,500	82,500	44,500	300	32,910,000
2010 Census	2,089,000	6,000	58,000	52,000	1,000	6,863,000
Other Pre2020	452,000	4,100	6,400	9,800	50	930,000
Legacy DSF						
Pre2020	906,000	200	43,000	32,500	250	33,280,000
Spring 2020	3,900	0	100	40	0	81,500
2020 Census Operation						
LUCA	283,000	2,100	20,500	14,500	60	5,641,000
CR	1,100	$N < 15$	0	0	0	149,000
NC	450	0	0	0	0	654,000
AdCan	1,700	$N < 15$	10,000	14,000	50	1,418,000
OBAV	0	0	0	0	0	892,000
FV	0	0	0	0	0	13,000
NRFU	0	0	0	0	0	200,000
UL	538,000	0	1,600	11,000	$N < 15$	616,000
UE/RA	0	7,400	0	0	0	7,400
GQ	0	0	102,000	60	4,900	107,000
TL	0	0	$N < 15$	174,000	$N < 15$	174,000
Total	7,372,000	41,500	324,000	353,000	6,600	83,930,000

NOTES: DSF, Delivery Sequence File; LUCA, Local Update of Census Addresses; CR, Count Review; NC, New Construction; AdCan, Address Canvassing; OBAV, Office-Based Address Verification; FV, Field Verification; NRFU, Nonresponse Followup; UL, Update Leave; UE/RA, Update Enumerate/Remote Alaska; GQ, Group Quarters; TL, Transitory Locations; SBE, Service-Based Enumeration.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0215.

5.5.2 Field/Clerical Operations

The four field/clerical operations added or touched over 50 million addresses, with the bulk of those coming from Address Canvassing. In the map analysis, we focus on the major address-building operations of OBAV and NRFU.

OBAV played an important role in the 2020 Census, adding almost 800,000 addresses alone and touching addresses from a variety of other sources. Figure 5.17 shows the geographic distribution by county while Figure 5.18 shows the distribution across tracts in Los Angeles, Orange, and San Diego Counties, California. Counties in eastern Kentucky, southern West Virginia, parts of the West, and municipios in Puerto Rico had decent fractions of their MAF units added or touched by OBAV. By design, OBAV followed up on

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

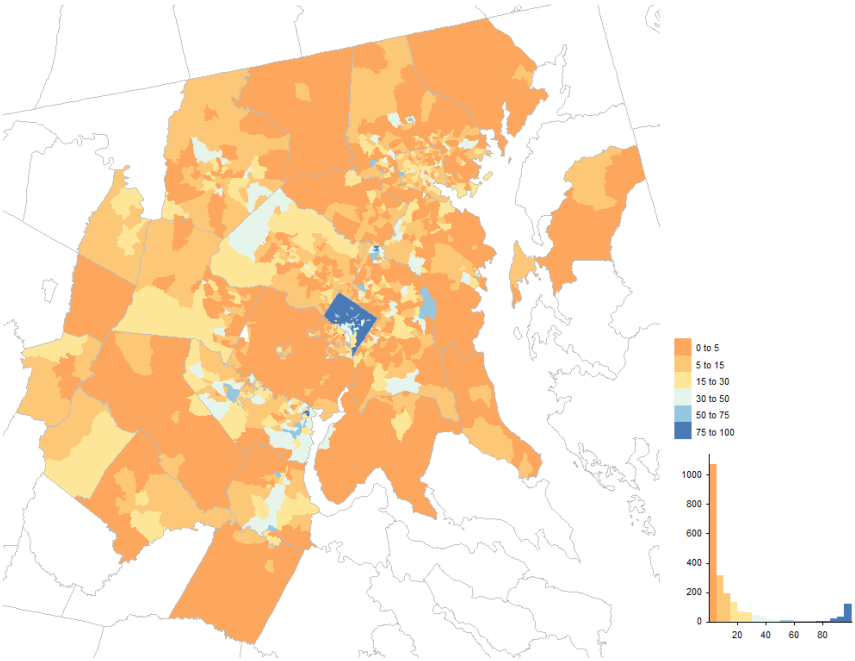


Figure 5.14 Percent of Master Address File units added or touched by 2020 Census Local Update of Census Addresses operation, by tract, Baltimore–Washington metropolitan area.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

addresses received from Non-ID Processing (i.e., Self-Response questionnaires without reference to a printed identification number from paper materials) that were not immediately matchable to the MAF. Hence, it is important to consider concentrations of OBAV returns in the context of concentrations of Non-ID Processing returns, which we derived from the dataset underlying the Census Bureau's Operational Quality Metrics releases and illustrate in Figures 5.19 and 5.20.

Similarly, NRFU interacted with addresses from almost all MAF source categories; see Figures 5.21 and 5.22 for the distribution by county and an illustration of tract-level variation in the state of Missouri. From a spatial perspective, counties in a belt stretching across the South from east Texas to South Carolina, along with counties in eastern Kentucky and southern West Virginia, had at least 9% of their MAF units added or touched by NRFU. There

Copyright National Academy of Sciences. All rights reserved.

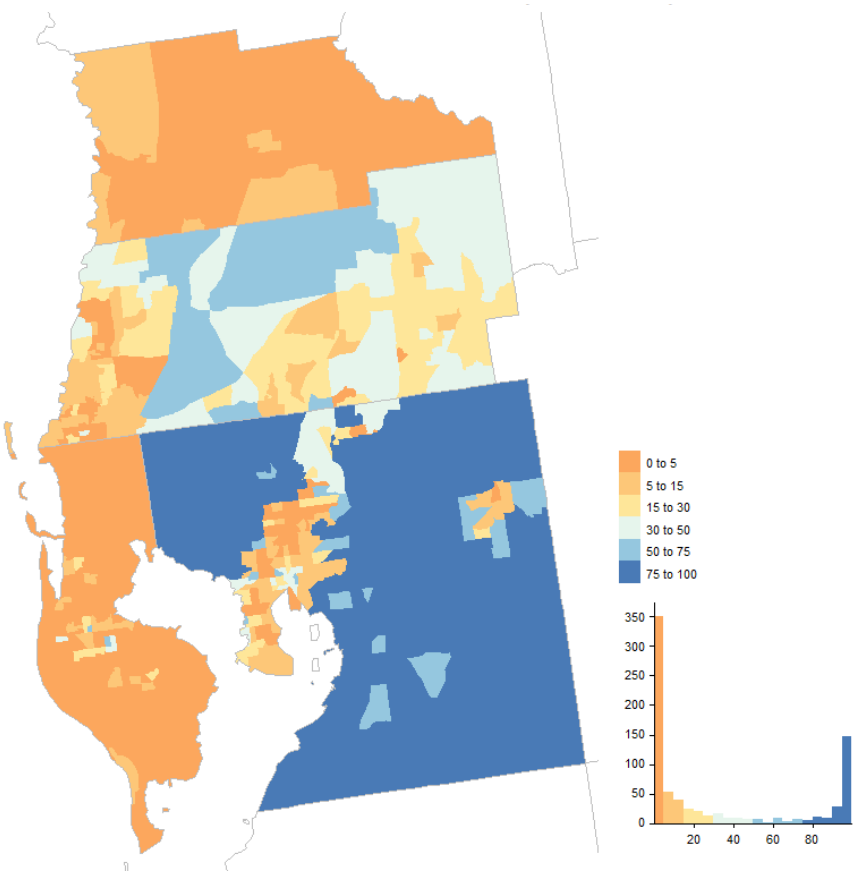


Figure 5.15 Percent of Master Address File units added or touched by 2020 Census Local Update of Census Addresses operation, by tract, Tampa–St. Petersburg metropolitan area.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

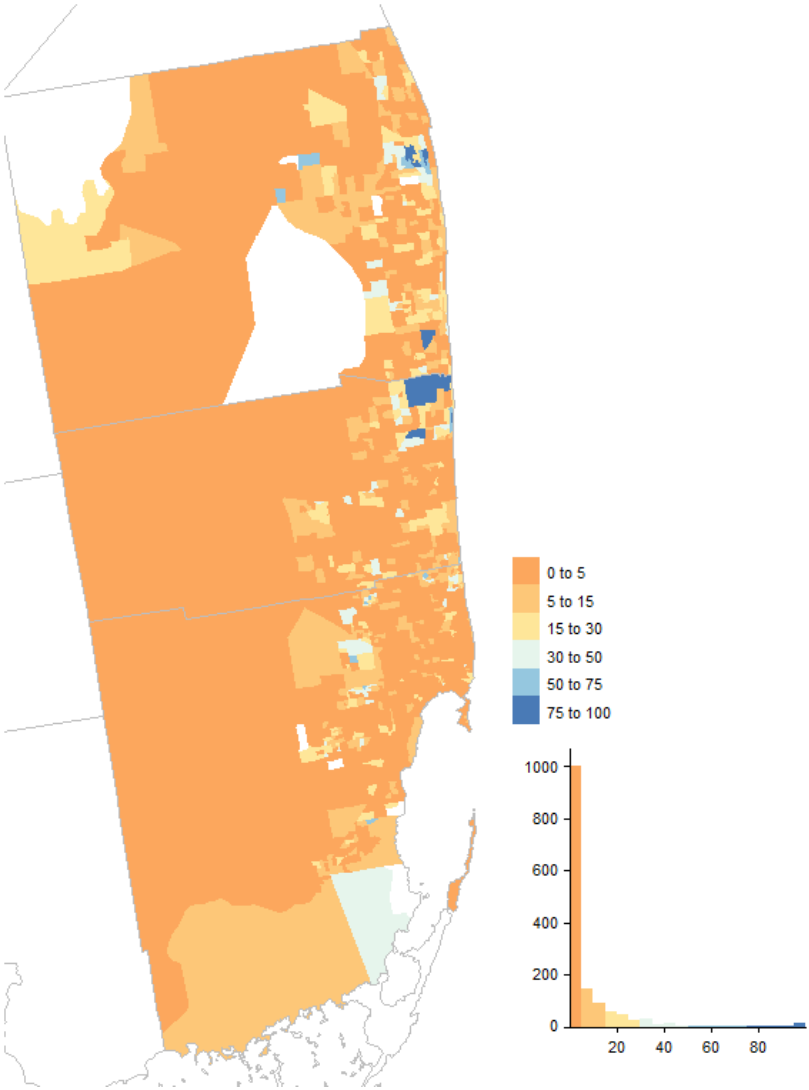


Figure 5.16 Percent of Master Address File units added or touched by 2020 Census Local Update of Census Addresses operation, by tract, Miami-Dade County metropolitan area.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

is also a belt of counties running from west Texas up to North Dakota and Montana where NRFU added or touched a sizable fraction of addresses.

5.5.3 Census Enumeration Operations

The five census enumeration operations added or touched approximately 8 million addresses, with the overwhelming majority (7.4 million) coming from UL. The other four operations targeted specific types of MAF entries (e.g., GQ or TL). UL areas were selected prior to the start of 2020 enumeration operations. West Virginia, southeastern Missouri, and counties stretching from west Texas up through Montana and Idaho had more than 50% of their MAF entries added or touched by UL (Figure 5.23). Even in heavily populated metropolitan areas such as San Bernardino-Riverside-Imperial, California, specific census tracts have a large fraction of addresses added or touched by UL (Figure 5.24).

Given the general prohibition on physical contact with households because of the COVID-19 pandemic, there were substantial delays in the execution of UL operations in areas designated for this method of enumeration. At the same time, the Census Bureau and its local partners encouraged households to self-respond, in the absence of a mail request and accompanying Census ID. While it is impossible to make any definitive connection, a comparison of Figures 5.23 (UL as a share of total MAF units) and 5.19 (Non-ID as a percent of Self-Response addresses) suggests that Non-ID was more common in UL areas. In both maps, concentrations appear in West Virginia into Kentucky, parts of the Great Plains, and a number of counties in western and southwestern states.

5.6 CONCLUSIONS

The most striking impression from analysis of MAF metadata is also the least surprising: the MAF is largely composed of addresses that have been on the rolls for many years, and continuity and sameness in the MAF over time necessarily dominates change and churn. This observation is decidedly not meant to diminish the importance of seeking new and updated address information; it actually serves to heighten the importance of both maintaining the existing MAF and finding ways to more efficiently update the list. Our analysis suggests the general findings that the Census Bureau's strategies for filtering the MAF were effective at identifying address records that were likely to be enumerable, and that the 2020 Census address-building operations made useful contributions, even though user-input operations like LUCA were not very efficient at producing unique information (yielding address records that did not result in enumeration in the 2020 Census or that were duplicative of already-known MAF entries). One particularly important 2020 Census operation that contributed to the MAF was OBAV, which primarily served to resolve addresses

Copyright National Academy of Sciences. All rights reserved.

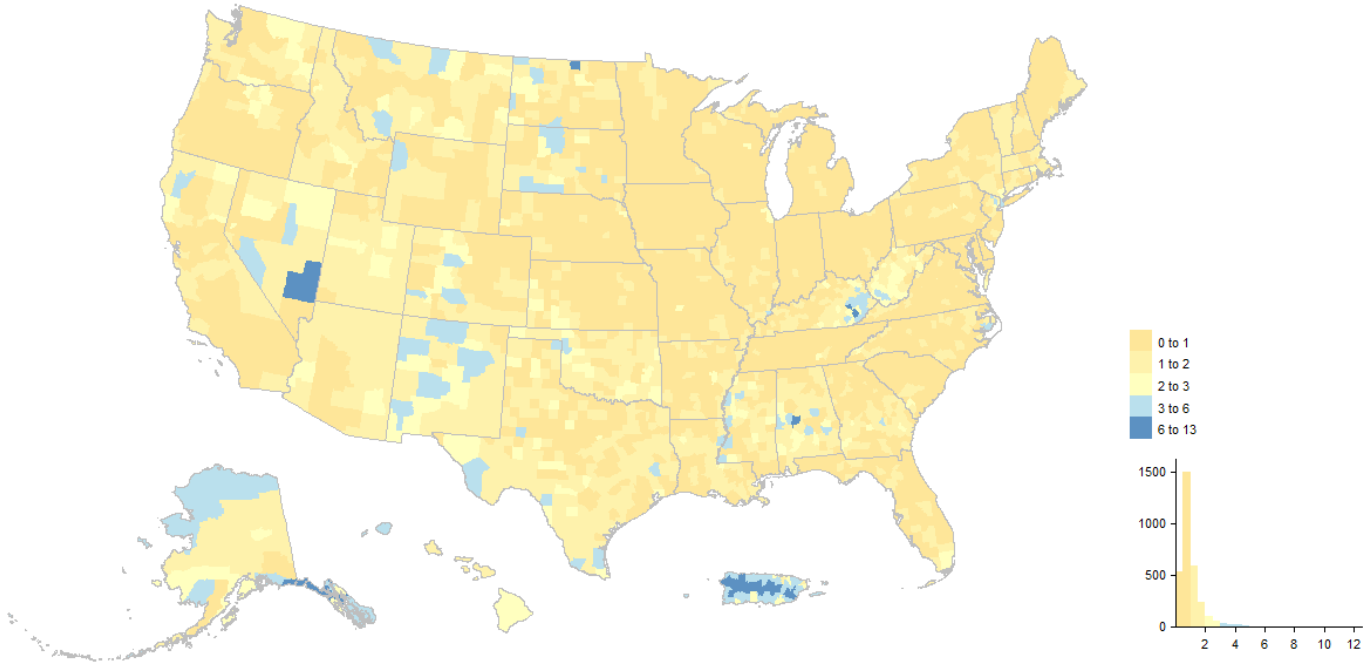


Figure 5.17 Percent of Master Address File units added or touched by 2020 Census Office-Based Address Verification operation, by county.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

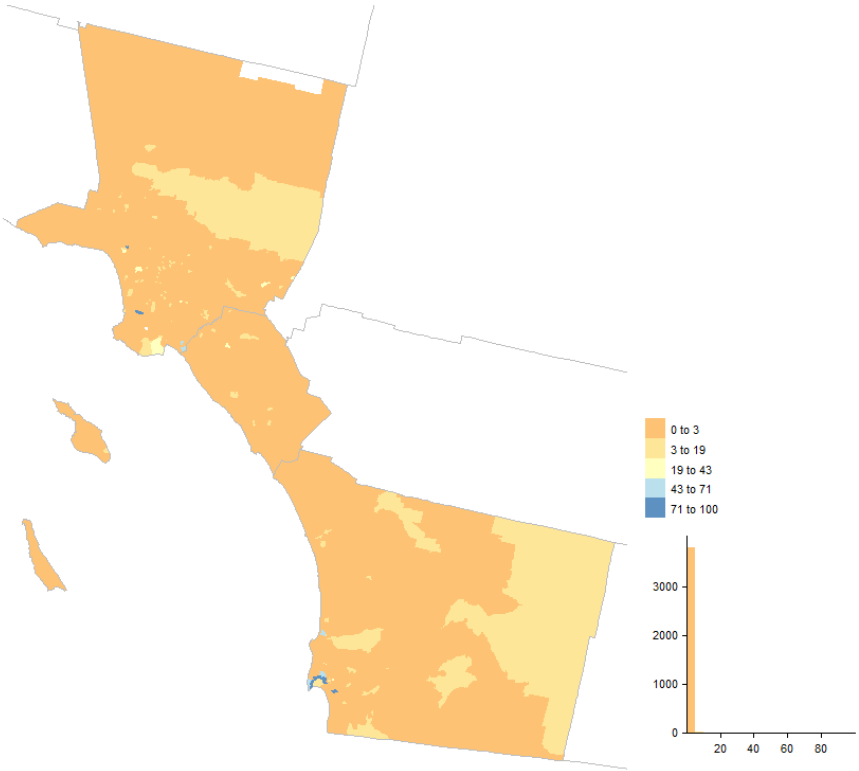


Figure 5.18 Percent of Master Address File units added or touched by 2020 Census Office-Based Address Verification operation, by tract, Los Angeles–Orange–San Diego Counties, California.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

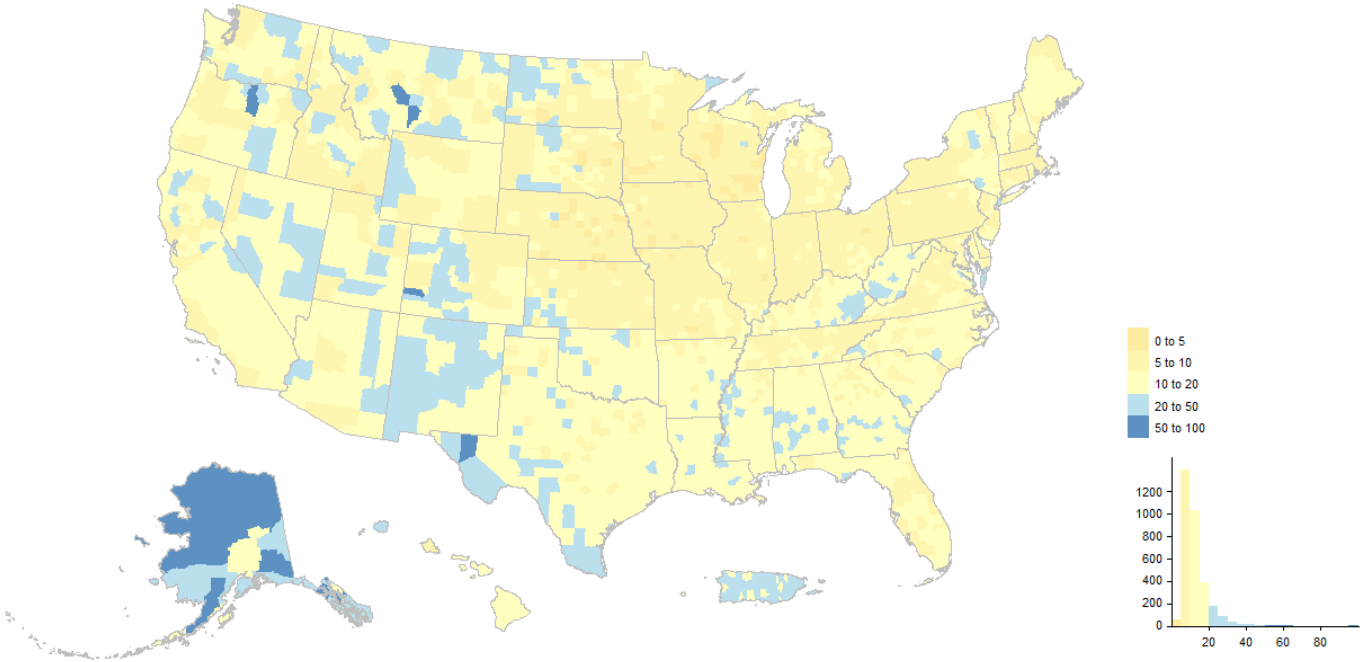


Figure 5.19 Percent of Self-Response housing units (occupied and vacant) using Non-ID Processing, by county.

SOURCE: Calculated from tract-level data underlying public 2020 Census Operational Quality Metric releases. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

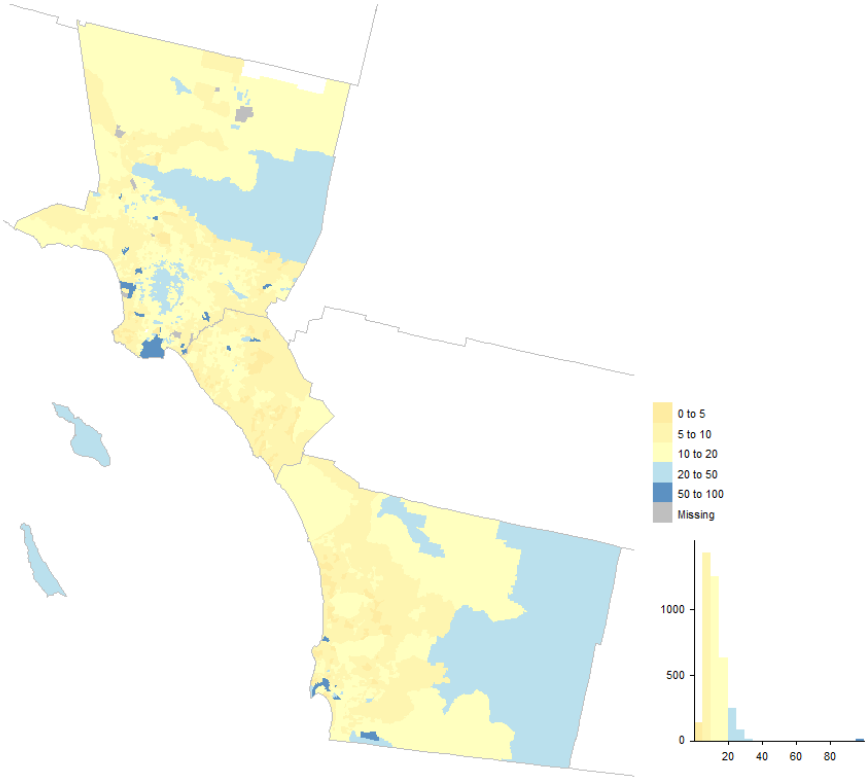


Figure 5.20 Percent of Self-Response housing units (occupied and vacant) using Non-ID Processing, by tract, Los Angeles–Orange–San Diego Counties, California.

SOURCE: Calculated from tract-level data underlying public 2020 Census Operational Quality Metric releases. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

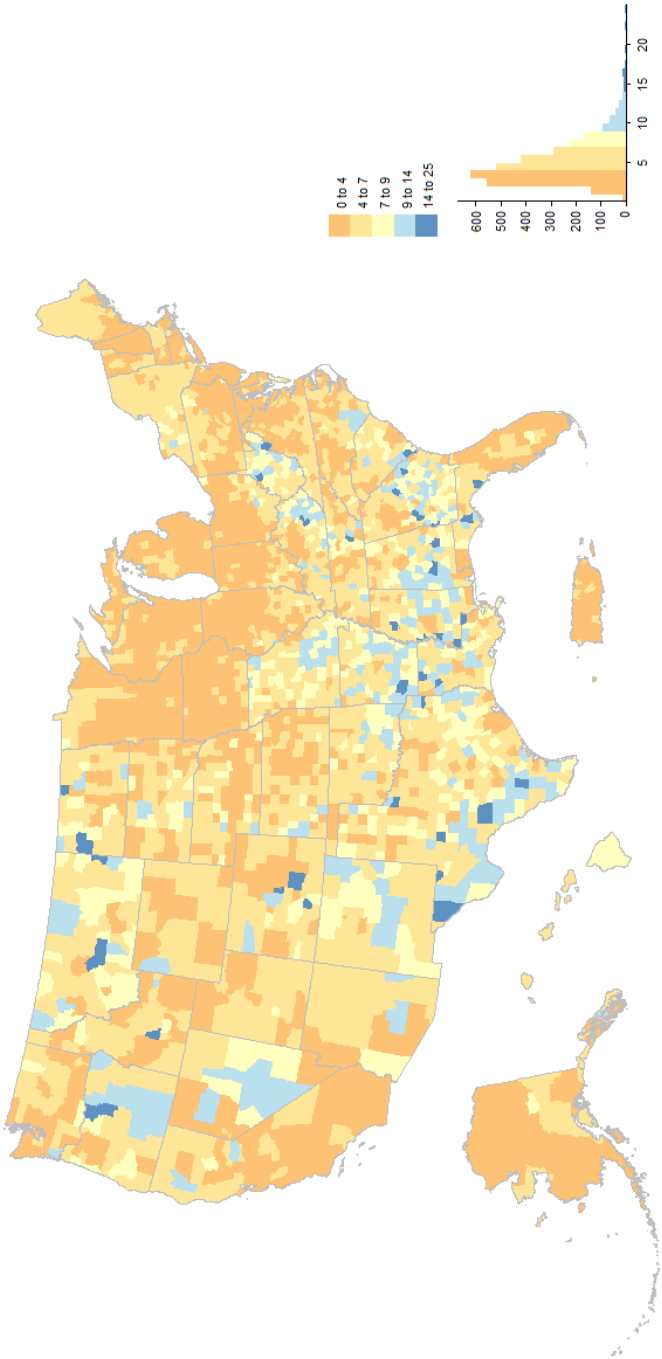


Figure 5.21 Percent of Master Address File units added or touched by 2020 Census Nonresponse Followup operation, by county.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

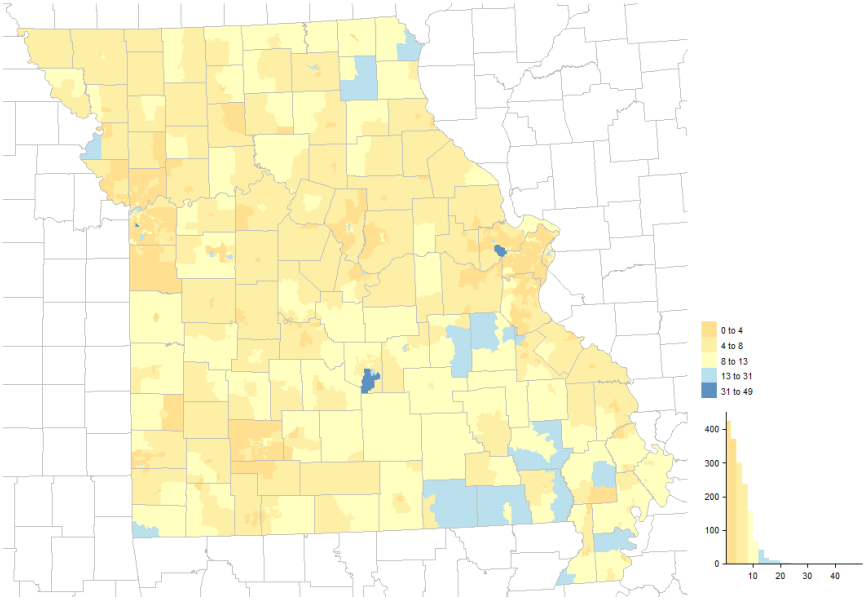


Figure 5.22 Percent of Master Address File units added or touched by 2020 Census Nonresponse Followup operation, by tract, Missouri.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

from Non-ID Processing (census returns without a preprinted MAFID code) that were not automatically matched to the MAF. In trying to resolve address information without field visits, and instead using other data and imagery resources, OBAV functioned much as a fully realized IOAC would have (with Active Block Resolution preserved as a component). The OBAV operation uniquely added roughly 750,000 new addresses to the MAF and appears to have been a major factor in "reactivating" many others (i.e., confirming that a respondent-supplied address matched an address already on the MAF but that had failed the filter to enter the starting base for the 2020 Census).

> *Conclusion 5.1: The U.S. Census Bureau's approach of retaining a fluid Master Address File database and filtering operational extracts from that database is a sound one, provided that the filters are continually studied and updated to improve their efficacy.*
>
> *Conclusion 5.2: The U.S. Census Bureau's reliance on the U.S. Postal Service's Delivery Sequence File as a primary update source between*

Copyright National Academy of Sciences. All rights reserved.

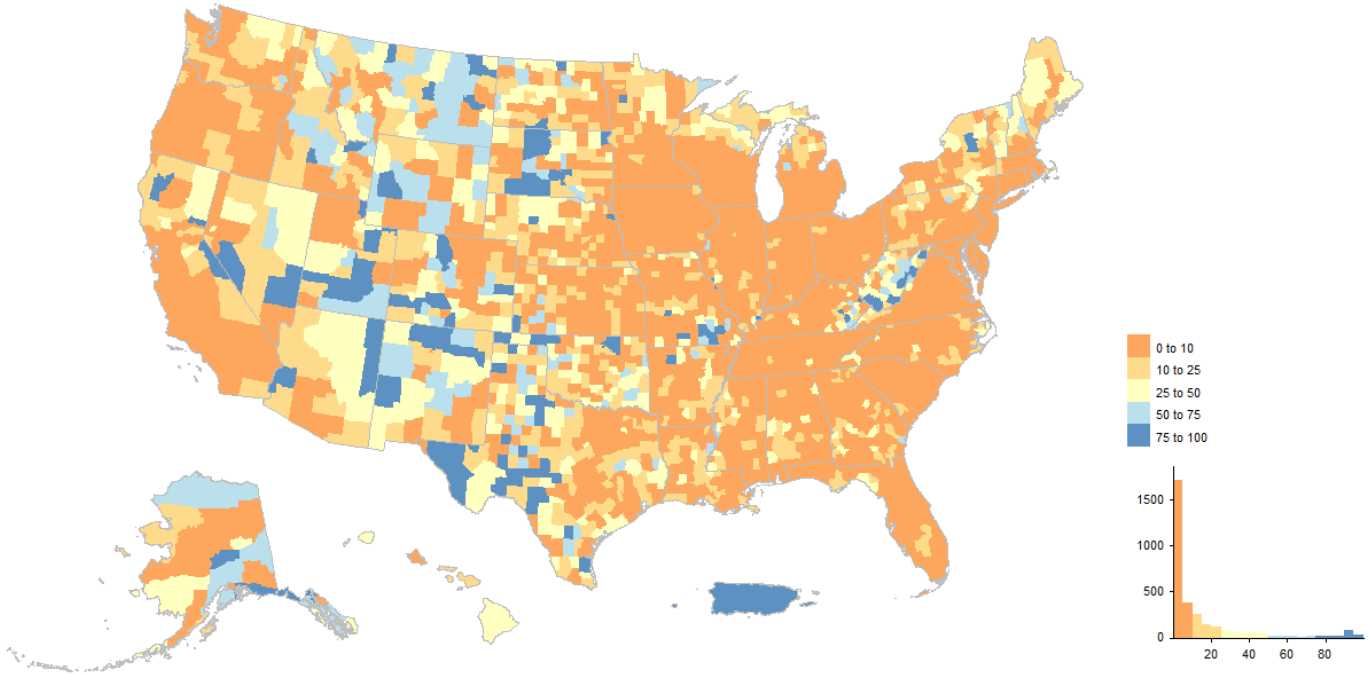


Figure 5.23 Percent of Master Address File units added or touched by 2020 Census Update Leave operation, by county.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

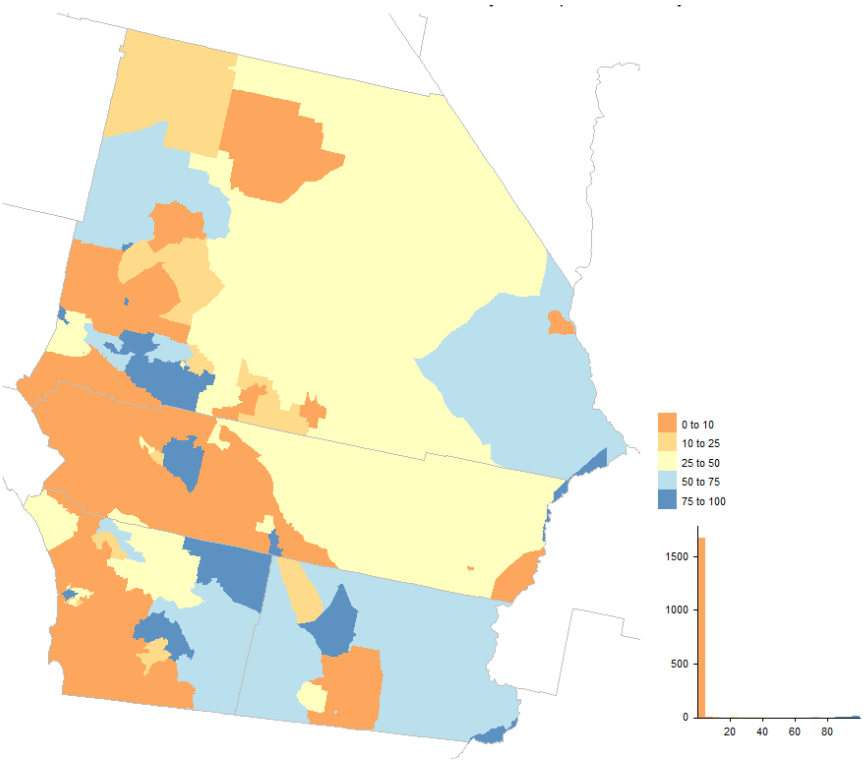


Figure 5.24 Percent of Master Address File units added or touched by 2020 Census Update Leave operation, by tract, San Bernardino–Riverside–San Diego–Imperial Counties, California.

SOURCE: MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0197.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

censuses is a good one, though it necessitates a reliable update source in areas without mail delivery.

Conclusion 5.3: With respect to the user-input address-building operations of Local Update of Census Addresses, Count Review, and New Construction, both participant governments and the U.S. Census Bureau erred on the side of inclusion—with local authorities submitting many more addresses for review than were actually in the operation's scope and the Census Bureau injecting many of those submitted addresses into the enumeration process to let fieldwork determine their accuracy. Consequently, these user-input operations appear to have added more bad (unenumerated) addresses to the 2020 Census than good (enumerated) addresses.

Conclusion 5.4: The Office-Based Address Verification operation for resolving respondent-supplied addresses from Non-ID Processing was an effective innovation in the 2020 Census, both in terms of its potential for resolving address listing issues with limited field work and for informing revisions to Master Address File filters.

Recommendation 5.1: The U.S. Census Bureau should continue to research and refine the filters it applies to the Master Address File to derive functional operational extracts, with the intent to reduce the number of addresses cancelled during collection (i.e., flagged as deletes in Nonresponse Followup and other field operations). Along those lines, the Census Bureau should continue to study ways to partner with and more fully utilize U.S. Postal Service data (including the Delivery Sequence File) and potential sources for address addition and revision information between censuses, such as address data recoverable from administrative records extracts, with particular eye toward more regular updating in areas without mail delivery.

Recommendation 5.2: The rules and scope of participant-input address-building operations of the census require clarification, including the Local Update of Census Addresses, New Construction, and Count Review. The U.S. Census Bureau should continue to improve ways to work with user-supplied input or data resources, and make it easier for state, local, and tribal authorities to supply input in usable form, while more clearly laying out the expectations for those address-building operations.

Recommendation 5.3: The U.S. Census Bureau should extend and improve during-the-decade programs, akin to the Geo-

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

graphic Support Systems Initiative of the 2010s, with the goal of continuous address- and map-database refinement. Such efforts should include engagement with state, local, and tribal partners, including on the question of whether some addresses can safely be removed from the Master Address File.

Recommendation 5.4: Lessons learned from the Office-Based Address Verification operation should be documented and incorporated into the tools and techniques for ongoing, continuous address and map data refinement (e.g., continuous In-Office Address Canvassing) as well as revisions of the Non-ID Processing operation.

Recommendation 5.5: The U.S. Census Bureau should research and test approaches to increase the proportion of dwellings that can be mailed to while still being associated with a physical geographical location. This would reduce the proportion of addresses in Update Leave and expand the use of approaches yielding higher Self-Response for more dwellings in the 2030 Census.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

– 6 –

Self-Response to the Census

A goal of every census is to count the entire population once, only once, and in the correct location. A related goal is to obtain complete and accurate data on the census form questions, such as race and ethnicity, age, and housing tenure (owner/renter), for people and households. The 1960 Census was the first to use self-response as the primary means of enumeration, based on research that showed wide variation in the quality of responses obtained by census enumerators (National Research Council, 2010:22–24). Evaluations of census quality in terms of coverage of the population and completeness of responses to individual questions have uniformly shown, including in 2020, that self-responses to the census are of higher quality than those obtained by interviewers in Nonresponse Followup (NRFU) and other in-person enumeration operations (see Box 6.1, which calls Figures 6.1 and 6.2, and Chapter 3).

The 2020 Census achieved a self-response rate—the percentage of all addresses that the U.S. Census Bureau endeavored to enumerate—that was roughly comparable to the 2010 Census mark, about 66–67%. The 2020 Census self-response return rate—the percentage of occupied housing units that self-responded—was about 74% according to information provided by the Census Bureau for the analysis in this chapter, examining the performance of the 2020 Census Self-Response options by characteristics of groups of census tracts.[1] The

[1]Comparisons of self-response and self-response return rates between 2020 and 2010 are not straightforward because the 2010 Census had a higher proportion of households that had no opportunity for self-response because they were enumerated by in-person procedures only (see Appendix C). We do not provide a self-response return rate for the 2010 Census because of this denominator problem and because rates that can be calculated for 2010 from the Census Bureau's

Copyright National Academy of Sciences. All rights reserved.

Box 6.1 Self-Responses Higher Quality than Enumerator Responses

Completeness of Coverage

While methods of estimating correct enumerations in the census by response mode (self-response versus enumerator response) are not strictly comparable, the 2000–2020 Censuses exhibit a clear pattern of self-responses having higher percentages of correct enumerations than enumerator responses (National Research Council, 2004a:Table D.1; Keller and Fox, 2012:Tables 11, 12; Hill et al., 2022:Appendix Tables 7, 8; see also Chapter 4). Similarly, the group of census tracts in the highest decile (10%) of self-return rates in both 2000 and 2020 had markedly lower rates of estimated omissions from the census than the group of census tracts in the lowest decile of self-return rates (Figure 6.1).

Completeness of Characteristics

Self-responses have historically had lower percentages of missing data for individual items, such as age, sex, and race/ethnicity, compared with enumerator responses. Similarly, self-responses compared with enumerator responses have historically had lower rates of whole-person imputations (people for whom no information was gathered in the census except that a person existed at an address). Figure 6.2 shows item imputation rates for age/date of birth, sex, race, and ethnicity in the 2020 Census for self-responses by type—internet, paper, telephone—and for enumerator responses by type—household member, proxy, administrative records (excluding such in-person operations as Update Enumerate and Remote Alaska). The differences among self-response types are minor, whereas the differences between self-response and enumerator response (especially those obtained from a proxy, such as a neighbor or landlord) are striking.

NOTE: In surveys with well-trained interviewers and sufficient time to obtain as much data as possible from respondents, data quality is often as good or better for interviewer-obtained responses as from self-responses. The reverse is true for the census because the enumerators receive minimal training, mostly have no or limited experience, have limited time to complete their work, and likely know that obtaining the population count at each household they visit is the priority—not necessarily the characteristics of individual household members. For data to this point, see National Research Council (2004a:Table 7.3), which compares self-responses and enumerator responses for the basic items (e.g., age, sex) on the 2000 Census long-form sample and the Census 2000 Supplementary Survey (the precursor to the American Community Survey).

Copyright National Academy of Sciences. All rights reserved.

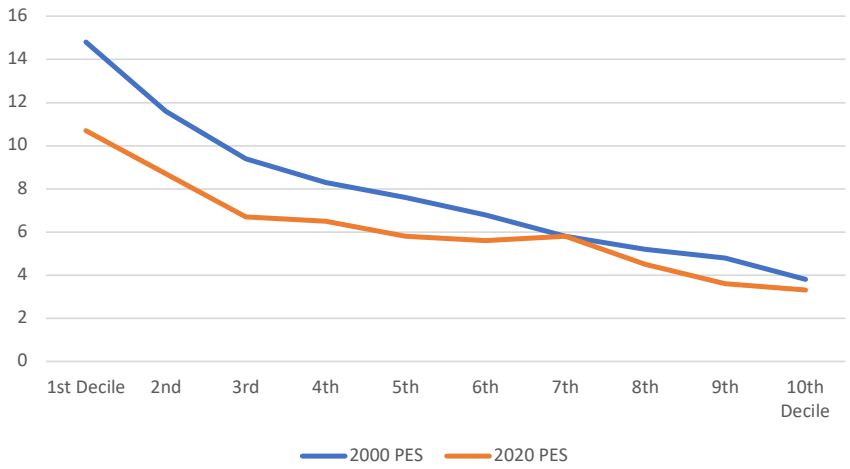


Figure 6.1 Percentages of people omitted from the census for the household population by self-return rate deciles for census tracts, 2000 and 2020 Census Post-Enumeration Surveys.

NOTES: Each decile contains 10% of the nation's census tracts, ordered by self-return rate from lowest (1st decile) to highest (10th decile). Update Enumerate areas, by design, are enumerated in person and hence fall into the first (lowest) decile of census tract self-return rates.

SOURCE: 2000: National Research Council (2004a:Table D.2). 2020: Hill et al. (2022:Appendix Table 6). A comparable analysis was not conducted for the 2010 Census.

solid self-response in 2020 occurred despite the COVID-19 pandemic, which curtailed local outreach and communication efforts, and despite low levels of public trust in government.[2] This achievement was vital to containing the workload for NRFU, which the COVID-19 pandemic made especially challenging to conduct.

In this chapter, we describe and assess the 2020 Self-Response modalities and outcomes. We first detail the Self-Response operations for 2020 (with a summary for 2010 as a frame of reference), highlighting the very successful

Operational Quality Metric releases (U.S. Census Bureau, 2021b:Address Resolution by Household Size) differ from official rates for that census (see Letourneau, 2012). We note that the 74% self-return rate for 2020 cited in the text also differs from rates that can be calculated from the same operational metrics. A forthcoming 2020 Census operational assessment report should provide definitive rates.

[2]See Pew Research Center (2022). The percentage of people agreeing that the federal government can be trusted to do the right thing "just about always" or "most of the time" has remained at or below 24% since 2007.

Copyright National Academy of Sciences. All rights reserved.

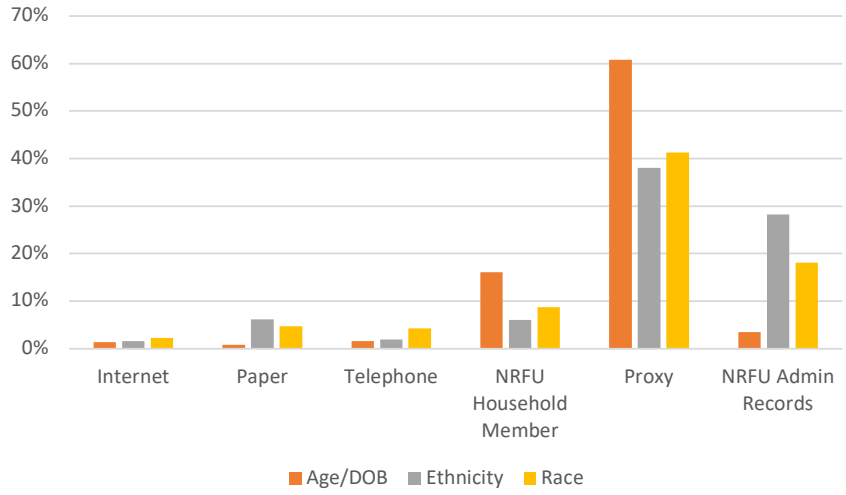


Figure 6.2 2020 Census percent of item nonresponse for age/date of birth, ethnicity, and race by Self-Response and Nonresponse Followup (NRFU) modalities.

NOTES: DOB, date of birth; NRFU, Nonresponse Followup.

SOURCE: Plotted from U.S. Census Bureau (2021c:Table 2).

introduction of the internet as a primary Self-Response mode in 2020 (see Appendix C for more details).

We then examine variation in self-response among groups of census tracts categorized by social, economic, and housing characteristics from the American Community Survey (ACS). Variation in self-response inevitably means variation in data quality among areas and groups. When measured by completeness of population coverage, variation in quality has adverse implications for the equity of allocations of U.S. representatives, funds, and services based on census results. Variation in completeness and accuracy of characteristics can also be detrimental to areas and groups to the extent that imputation procedures for item missingness and edits to correct inconsistent responses may also be inaccurate. Both kinds of variation can erroneously reduce or amplify true differences among geographic areas and population groups and thereby bias comparisons of many kinds of rates that use census data for denominators—for example, mortality rates by age, sex, race, and ethnicity in counties or smaller areas.

Copyright National Academy of Sciences. All rights reserved.

Another issue with variations in completeness of coverage and accuracy of characteristics is that, for the balance of the decade, decennial census data become the statistical foundation for the Census Bureau's annual population estimates for states, counties, cities, and towns. In turn, the population estimates have many uses, including as denominators and as weighting controls for the ACS and other surveys. Variations in data quality among geographic areas and population groups can undermine the relative accuracy of population estimates and denominators from the census.[3]

6.1 SELF-RESPONSE STRATEGIES AND PROCESSES

Except for a small-scale experiment with internet response in the 2000 Census, the primary Self-Response mode in the 1970–2010 Censuses continued to be paper questionnaires as in 1960, with a small number of self-responses obtained by telephone. Peer countries began offering the internet as a self-response mode as early as the 2001 census in Spain and the 2006 census in Canada and several other countries and expanded its use substantially in censuses in 2010–2011 (National Research Council, 2010:290–301). The Census Bureau benefited from their experience when it made the decision to support Internet Self-Response for the 2020 Census, as recommended by the National Research Council (2011:1,18–20). The Census Bureau also benefited from its own experience in successfully using the internet as a primary self-response mode in the ACS, beginning in 2013. Moving to operationalize and promote Internet Self-Response in 2020 proved to be one of the major innovations and success stories of this census. Internet Self-Response was crucial to enabling the Census Bureau to surmount the hurdles for the census enumeration posed by the COVID-19 pandemic. Responses with a paper questionnaire were also important in achieving higher-than-2010 self-response and return rates.

6.1.1 2010 Census Self-Response Procedures

In 2010, bar-coded paper questionnaires with a return envelope for self-response were mailed to addresses in areas of the country with predominantly "city-style" addresses (street names and house numbers). These Mailout/Mailback areas comprised type of enumeration area (TEA) 1, which accounted for 91% of 115.9 million total addresses. Another 8% of addresses in

[3]The Census Bureau's population estimates currently use only population totals from the 2020 Census. For demographic characteristics, they use 2010-based estimates trended forwarded. The reason has to do with the new Disclosure Avoidance System (DAS), using differential privacy-based methods. The 2020 DAS has precluded the availability of the necessary demographic data for the population estimates until an (as of fall 2023) unspecified date. See Chapter 11 for a detailed discussion of the new confidentiality-protection methods and their implications for data quality for small geographic areas and population groups.

Copyright National Academy of Sciences. All rights reserved.

TEA 2 (Update/Leave areas) and TEA 7 (Urban Update/Leave areas) had the questionnaire dropped off by a census employee with a return envelope for self-response. The remaining 1% of addresses were enumerated in person (Johanson et al., 2011:15).

The mailing strategy in TEA 1 consisted of three mailings to every address: (1) an advance letter informing households that the census questionnaire would be arriving soon; (2) a mailing package containing the questionnaire; and (3) a reminder postcard. A fourth mailing, which contained a replacement questionnaire, was sent to 15 million addresses in Mailout/Mailback areas expected to have low response and another 25 million nonresponding addresses in Mailout/Mailback areas expected to have a moderate level of self-response.

The Census Bureau also included a "Be Counted" operation, which encouraged people to pick up and complete an unaddressed ("Non-ID") questionnaire at a library or other public location. This procedure added the people in 676,000 housing units to the 2010 Census (addresses provided by respondents had to match a housing unit address on the Master Address File (MAF) to be counted—see Jackson et al. (2012:Table 27)).

6.1.2 2020 Census Self-Response Procedures

During the 2010–2020 intercensal period, the Census Bureau made a commitment to heavily promote Internet Self-Response in 2020, to contain costs and improve data quality. The Census Bureau tested a variety of strategies to improve its mailout procedures, to develop a user-friendly online questionnaire and log-in page, and to best ensure that the cloud-based internet operation would not crash under heavy loads or be attacked; Box 6.2 provides additional comment on this internet response platform. The Census Bureau also developed procedures for "Non-ID" online response whereby people who had mislaid their 12-digit internet ID codes or had not received them could provide their addresses and complete their enumerations on their cell phones or laptops.

Testing resulted in generation of a more elaborate "wave" mailing strategy (pioneered in Canada) to encourage self-response for addresses in TEA 1. In 2020, TEA 1 comprised 95% of the total of 151.8 million addresses, while TEA 6 (Update Leave), in which census enumerators delivered paper questionnaires accompanied by letters with an internet ID code to respond online, comprised another almost 5%. The remaining TEAs involving in-person enumeration (Update Enumerate and Remote Alaska) comprised less than 1% of addresses.[4] Advertising and outreach activities supported all of these enumeration strategies.

[4]See https://www.census.gov/newsroom/press-releases/2019/tea-viewer.html.

Copyright National Academy of Sciences. All rights reserved.

Box 6.2 Primus for Internet Response

Discussion of self-response in the 2020 Census is an appropriate moment to note and commend two pieces of in-house, within-the-Census-Bureau systems development that were instrumental in facilitating that self-response. The U.S. Census Bureau's Integrated Computer Assisted Data Entry (iCADE) digital-imaging and data-capture system was developed over several decades and, with a modernized codebase in the mid-2010s, was tapped to handle all of the capture-from-paper work in the 2020 Census. More importantly, a home-grown system proved capable of handling the load of Internet Self-Response as the primary response channel to the 2020 Census, weathering a very last-minute call into service to do so.

In the early and mid-2010s, Census Bureau research and development teams did considerable work in fielding prototype data-collection systems and control systems to manage the enumeration, as part of work on the key innovation areas for the 2020 Census. One such thread in this work was upgrading the response system then in use for internet surveys at the Census Bureau, also an in-house system known as Centurion. The new system, Primus, was designed to nimbly allow for a heavy concurrent volume of users, while also being easy to use and accessible.

The Census Bureau was then engaged in the Census Enterprise Data Collection and Processing (CEDCaP) initiative, working on information technology solutions for the full sweep of Census Bureau operations. In May 2016, the Census Bureau cited perceived benefits in safety and scalability in contracting with Pegasystems Inc. to provide a commercial off-the-shelf software base for its technical systems, including many of those intended for use in the 2020 Census (U.S. Census Bureau, 2016b), attempting to balance these benefits with the incurred risk of time and effort in recoding existing systems. Fortunately, Primus was not entirely discarded in this effort; rather, it was to continue in development on a secondary, back-up basis. The exact details of subsequent developments, outside the public eye, are unclear, but conversion to the Pega base proved more difficult than expected and concerns about systems readiness and vulnerability became prominent as the 2020 Census drew close. In particular, reporting in December 2019 suggested that the Census Bureau had been positioning itself "to build Primus out, in case it was needed," since summer 2019. A related U.S. Government Accountability Office (2020) report became the forum for revealing the February 7, 2020, announcement by the Census Bureau that other 2020 Census systems would continue to use the Pega base but that Primus would replace the Pega-based Enterprise Censuses and Surveys Enabling Platform–Internet Self Response (ECaSE-ISR). The tipping point was said to be scalability; the Census Bureau had set a goal of enabling up to 600,000 concurrent internet users (five times its expected concurrent user load), and Primus was able to accommodate that volume without performance issues while ECaSE-ISR was not. Shifting to the backup role, ECaSE-ISR was slated for continued use in the 2020 Census Telephone Questionnaire Assistance operation.

A change in systems less than a month before public roll-out was certainly not without risk, particularly because Primus had not been used in the last major precensus test in 2018. But, to the credit of the Primus development staff, the last-minute change was very successful. Summaries of the progress of the 2020 Census—amidst all the challenges—would take pride in the fact that the Primus response platform experienced no downtime and maintained cybersecurity throughout the period when internet responses were accepted for the 2020 Census.

Copyright National Academy of Sciences. All rights reserved.

Internet First and Internet Choice

For TEA 1, the first part of the mail strategy was to divide areas into "Internet First" and "Internet Choice." Internet First areas were those for which the majority of residences had relatively easy access to the internet—specifically, census tracts not characterized as Internet Choice (see below). For Internet First areas, the plan was to start with a letter from the Census Bureau, followed by a second letter, then a postcard, followed by a letter with a paper questionnaire included, and lastly an "It's not too late" postcard. Each letter informed residents about how to be enumerated online and provided them with a 12-digit ID to use for internet response.

Internet Choice areas were census tracts with low overall self-response rates in the ACS and one of the following: low ACS internet response rates, high rates of people ages 65 and older, or low rates of internet subscribership, estimated from 5-year ACS data for 2013 through 2017. The sequence of mailings for Internet Choice areas was to be a letter with a questionnaire enclosed along with a link for responding online, followed by a letter, a postcard, a letter with a questionnaire, and then an "It's not too late" postcard.

The first three mailings for both Internet First and Internet Choice areas went out on schedule beginning in mid-March 2020. The COVID-19 pandemic led to delays of 1–3 weeks for the remaining two mailings because of the need to protect Census Bureau workers who assembled mailing packages (see Table C.2 in Appendix C). In response to the pandemic, the Census Bureau added a sixth mailing of a postcard with an internet link sent in late July to nonresponding households in Internet First and Internet Choice areas. The Census Bureau sent a seventh mailing of a letter with a paper questionnaire in late August–early September to nonresponding Internet First households in low-response census tracts. In addition, the Census Bureau sent a postcard to 4.5 million U.S. Postal Service (USPS) boxes in areas where there was no USPS carrier delivery service. Finally, the Census Bureau extended the deadline for self-response from July 31, 2020, to October 15, 2020.

Non-ID Response

In 2020, the Census Bureau widely advertised the Non-ID response option via the internet (people could also respond without an ID by telephone), with considerable success. Before accepting a Non-ID response, the Census Bureau had to verify the address. Of Non-ID responses submitted on behalf of fully 17.6 million households, only 3.6% had to be discarded because the address did not match the MAF and could not be verified in the field (see Table C.3). In all, Non-ID responses constituted 9.4% of household enumerations, compared

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

with 55.6% for internet and telephone ID responses and 11.6% for paper-based responses (see Table 3.4).[5]

Evaluation of Self-Response Strategies and Processes

With the available data, we were unable to analyze the 2020 Census mailing and contact strategies beyond gross effects, such as somewhat earlier elevated levels of paper-questionnaire take-up in Internet Choice areas. We also could not disentangle the effects of messaging and mail strategies (e.g., Internet First and Internet Choice) in total or by geographic areas or population groups, nor how response mode choices were affected by the COVID-19 pandemic.[6] These are areas in which the Census Bureau could usefully mine its 2020 data to plan informative tests and experiments to determine optimal self-response modes and enumeration operations for the 2030 Census.

6.2 ANALYSIS OF VARIATION IN SELF-RESPONSE

Chapter 3 provided evidence of the overall superior quality of Self-Response—whether measured by population coverage, item and whole-person nonresponse, or age heaping—compared with responses obtained by enumerators. This chapter focuses on variations in self-response among small areas (census tracts, which average about 4,000 people) grouped by characteristics such as income and availability of broadband. Variations have implications for equity for many uses of census data because areas with low self-response will, other things equal, have lesser-quality data than other areas.

6.2.1 Previous Work on Self-Response for Census Tracts

Previous research on self-response for census tracts includes a report of the 2020 Census Quality Indicators Task Force (2020:App. 2), O'Hare (2021), and O'Hare and Lee (2021). These reports all used publicly available self-response rates for which the denominator included not only occupied housing units but also vacant units and addresses that were found not to be residential or not to exist. This denominator can confound comparisons between areas that differ in vacancy rates, for example, but the results of the three aforementioned studies are nonetheless informative.

[5]Little information is available on the quality of Non-ID self-responses compared with ID responses. One analysis from the 2020 Post-Enumeration Survey found that Non-ID self-responses were less likely to be correct enumerations and more likely to include whole-person imputations than ID self-responses, whether by internet, telephone, or paper, and than NRFU enumerations with the household head. Non-ID responses were better in these regards than NRFU enumerations with proxies and other household members (see Table 3.4).

[6]Chapter 10 has limited data on response mode choices by race and ethnicity.

Copyright National Academy of Sciences. All rights reserved.

The 2020 Census Quality Indicators Task Force (2020) compared 2010 and 2020 self-response rates for census tracts, counties, and states. Focusing on census tracts, the 2010 and 2020 Census rates were highly correlated, but the 2020 distribution of rates had more tracts at the high and low ends of self-response compared with 2010. The Task Force also provided evidence that census tracts with low internet response ended up with lower overall self-response—in other words, paper and telephone did not compensate in these tracts.

O'Hare (2021) found a high correlation between self-response rates for the nation's approximately 84,000 census tracts in the 2010 and 2020 Censuses, but also found that 4,000 tracts (about 5%) had an increase in self-response of 10 percentage points or more, and 8,000 tracts (about 10%) had a decrease in self-response of this magnitude between 2010 and 2020. Using ACS 5-year data to identify majority populations in each census tract, O'Hare determined that low-income tracts had the highest percentage with a decrease in response of 10 percentage points or more (23%), followed by Hispanic tracts (20%), low-education tracts (19%), Black Alone or in Combination tracts (15%), and immigrant or renter tracts (14% each).

O'Hare and Lee (2021:10–14) summarized their method of analyzing variations in self-response rates for census tracts. They identified historically undercounted population groups from research on prior censuses and identified census tracts with high concentrations of each group (defined as 50% for some groups and 33% for others). They excluded census tracts in Remote Alaska and Update Leave areas. O'Hare and Lee (2021:5) also found high correlations of 2010 and 2020 self-response rates and relatively low self-response rates in 2020 in census tracts with concentrations of Black people, Hispanic people, American Indian or Alaska Native people, foreign-born people, people with low incomes, people with a high school diploma or less education, people who rent, people who recently moved, and people who live in crowded housing. Our analysis of self-return rates (self-response as a percentage of occupied housing units) has similar results but shows more detail by mode of self-response and by examining the range of concentration of each group across all census tracts divided into ventiles (groups accounting for five percent of the total).

6.2.2 Panel's Analysis of Self-Return Rates

The Census Bureau provided the panel with self-response data from the 2020 Census operation—specifically with self-return rates for census tracts where the denominator is occupied housing units, not all addresses in the census. The operational data used (e.g., self-response completions by time) are not publicly available. Only operational metrics and ACS public data were used. Data for census tract percent of occupied housing units responding by Self-Response (internet, paper, telephone), and census tract 2016–2020 5-year ACS data from

Copyright National Academy of Sciences. All rights reserved.

Demographic Profiles were merged. These data were used to calculate zero-order bivariate correlation coefficients between each variable in the merged dataset. The correlation coefficients between each modality and each ACS variable were graphed from high to low for each modality.

For each ACS variable, the values were used to rank census tracts, and they were then grouped from highest to lowest into up to 20 categories (ventiles). Census tracts with missing data for an ACS variable were dropped. The means of each ACS variable were calculated for each ventile category, as was the mean of the percentage of households enumerated using each self-response mode. The resulting data are then employed to generate line graphs illustrating the variation in modality from one ventile to the next and across all ventiles.

Self-Response Correlations with Census Tract Social and Economic Characteristics

Zero-order bivariate correlations provide some indication of the association between two variables. They do not control for confounding factors. Figures 6.3–6.6 illustrate correlation coefficients between social, economic, and housing characteristics for total Self-Response, internet, paper, and telephone, respectively. Each plot is sorted from highest to lowest coefficient. Total Self-Response is positively associated with higher socioeconomic status (SES) indicators (median rooms, income, owner-occupied units, and non-Hispanic White), and negatively associated with lower SES indicators (percent of renter units, percent of vacant housing units, percent Hispanic, unemployment rate). While percent having broadband was the strongest association with Internet Self-Response, internet response looks similar to the total, implying that Internet Self-Response was driving total Self-Response.

Thus, those census tracts with a high percentage of households with broadband are likely to be represented more accurately in 2020 Census data than those tracts with small percentages of households with broadband. This sort of pattern, in which Internet Self-Response is positively associated with an ACS variable, suggests better-quality data for the populations living in those census tracts with greater prevalence of the characteristic represented by the variable.

Most households completed the census by internet, and paper was the second-most-common modality. Correlations by SES appear to be the opposite for paper compared with those we saw with Internet Self-Response, with lower SES characteristics positively associated with paper response. Given that paper has the highest item nonresponse compared with other self-response modalities, those census tracts with higher percentages of lower SES indicators are more likely to be represented in the 2020 Census with poorer-quality data.

Telephone Self-Response accounted for a small percentage of the total enumeration, yet it was positively associated with indicators of lower SES.

Copyright National Academy of Sciences. All rights reserved.

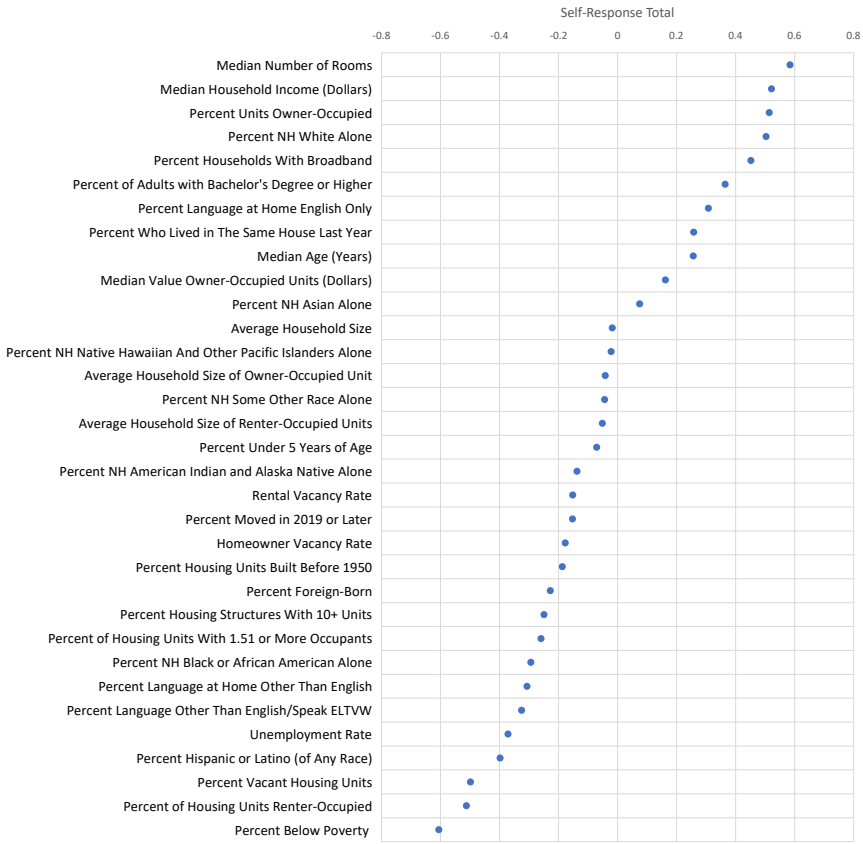


Figure 6.3 Zero-order correlations between total Self-Response return rate and American Community Survey socioeconomic characteristics, census tracts.

NOTES: NH, non-Hispanic; ELTVW, English less than very well.

SOURCE: 2020 Census Self-Response Return Rate Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0206.

Copyright National Academy of Sciences. All rights reserved.

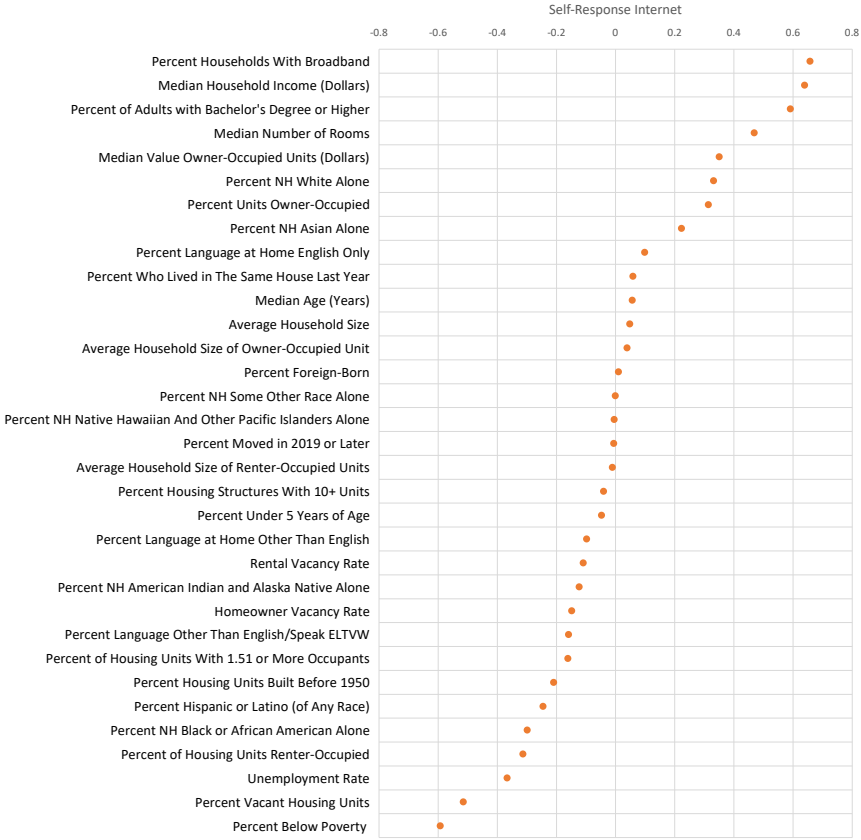


Figure 6.4 Zero-order correlations between Internet Self-Response return rate and American Community Survey socioeconomic characteristics, census tracts.

NOTES: NH, non-Hispanic; ELTVW, English less than very well.

SOURCE: 2020 Census Self-Response Return Rate Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0206.

Copyright National Academy of Sciences. All rights reserved.

Figure 6.5 Zero-order correlations between Paper Self-Response return rate and American Community Survey socioeconomic characteristics, census tracts.

NOTES: NH, non-Hispanic; ELTVW, English less than very well.

SOURCE: 2020 Census Self-Response Return Rate Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0206.

Copyright National Academy of Sciences. All rights reserved.

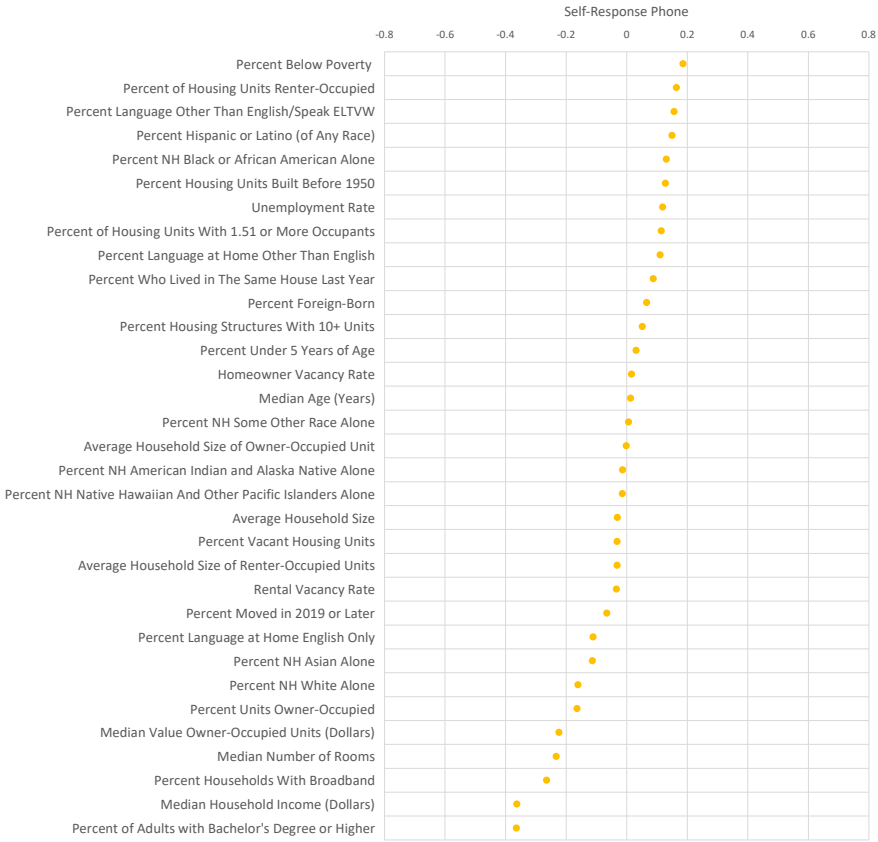


Figure 6.6 Zero-order correlations between Telephone Self-Response return rate and American Community Survey socioeconomic characteristics, census tracts.

NOTES: NH, non-Hispanic; ELTVW, English less than very well.

SOURCE: 2020 Census Self-Response Return Rate Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0206.

Copyright National Academy of Sciences. All rights reserved.

Given lower levels of item nonresponse for telephone compared with paper, telephone appears to slightly offset the quality issue for lower SES indicators and lower quality associated with paper.

Self-Response Ventile Plots

The percentage of enumerations completed by Self-Response mode is illustrated in Figures 6.7–6.12. Plots were generated for a select set of ACS variables illustrating the percent enumerated by each mode of Self-Response for each ventile of the ACS variable. Figure 6.7(a) illustrates household size. Larger and smaller households appear less likely to self-respond compared with those in the 2.5–3 persons-per-household range. Figure 6.7(b) illustrates the percentage of adults with bachelor's degrees or higher. Those census tracts with a small percentage of people with bachelor's degrees or higher appear to be much more likely to respond by paper compared with internet. Thus, census tracts with lower educational attainment are more likely to have poorer-quality 2020 Census data than census tracts with higher educational attainment. Figure 6.8(a) illustrates median household income, and the patterns by modality look very similar to those for percent with a bachelor's degree or higher.

Figures 6.9–6.10 focus on basic race and ethnicity groupings as well as linguistic isolation. We see lower Self-Response and Internet Self-Response for areas where smaller percentages of population consisted of non-Hispanic White people. Patterns for percent African American people, Hispanic people, and people who speak a language other than English or speak English less than very well appear to be similar to each other, with each having lower Self-Response and Internet Self-Response, and higher Paper Self-Response. Areas with high percentages of non-Hispanic White people and areas with low percentages of African-American people, Hispanic people, and people who speak a language other than English appear to be less likely to respond by internet and more likely to respond by paper. We mapped these ventiles and the census tracts appeared mostly in rural areas. Thus, the Census Bureau's strategies for Update Leave and Internet Choice, and perhaps limited internet access may result in poorer-quality data in these areas.

Figures 6.11–6.12 focus on housing stock characteristics. The plots demonstrate that areas with higher percentages of rental housing, vacant housing, and large housing-complex housing all have lower Self-Response and Internet Self-Response; however, they also have lower Paper Self-Response, resulting in a higher percentage of units in those census tracts being enumerated with NRFU. Areas with housing-related characteristics associated with lower SES are likely to be presented in the 2020 Census with poorer-quality data.

Copyright National Academy of Sciences. All rights reserved.

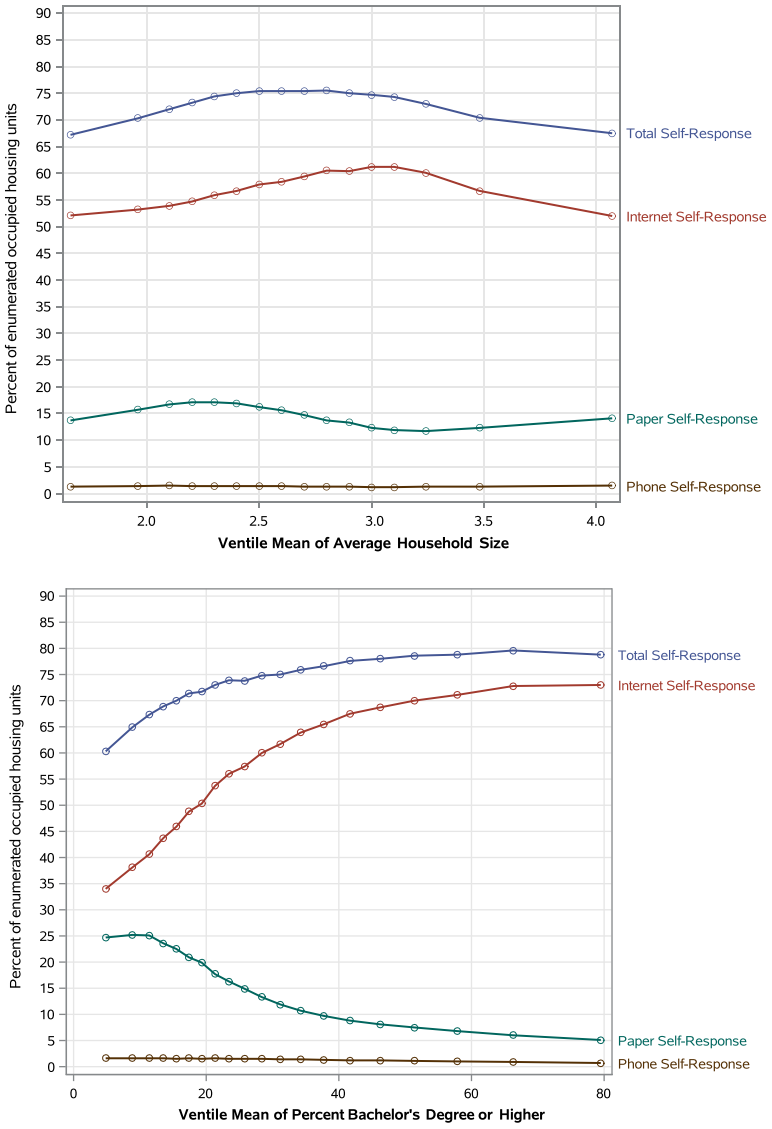


Figure 6.7 Mean Self-Response return rate by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) average household size and (b) percent bachelor's degree or higher.

SOURCE: 2020 Census Self-Response Return Rate Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0224.

Copyright National Academy of Sciences. All rights reserved.

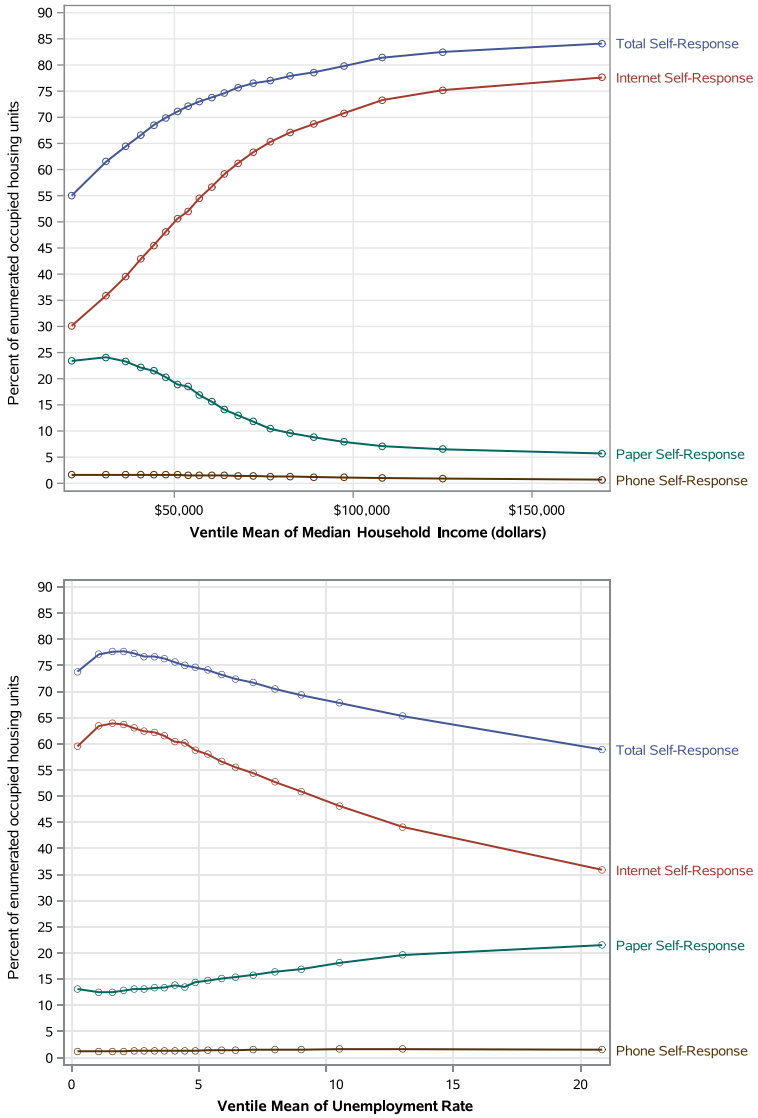
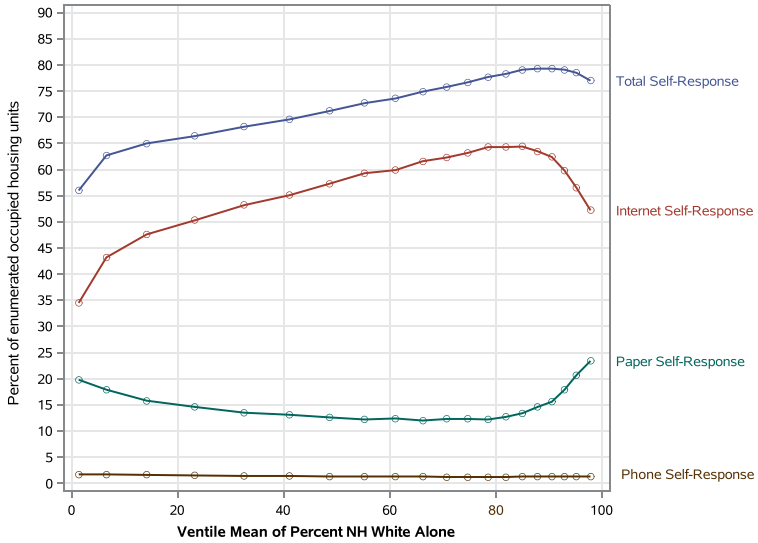


Figure 6.8 Mean Self-Response return rate by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) median household income and (b) unemployment rate.

SOURCE: 2020 Census Self-Response Return Rate Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0224.

Copyright National Academy of Sciences. All rights reserved.



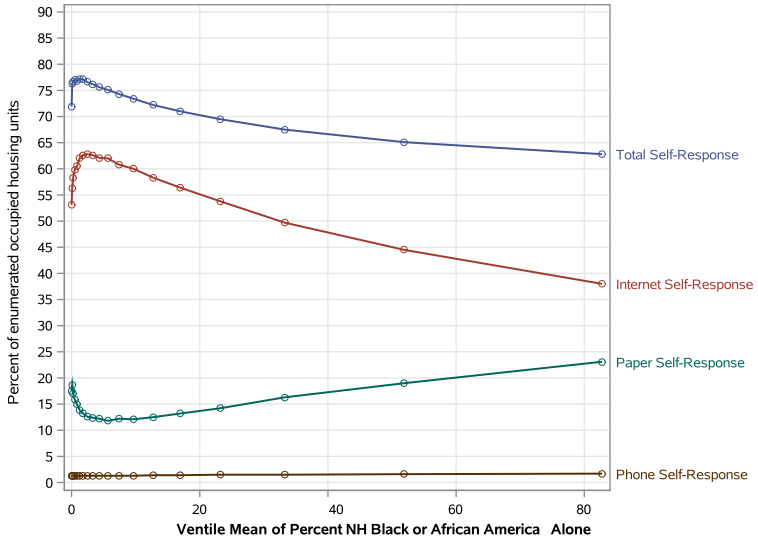
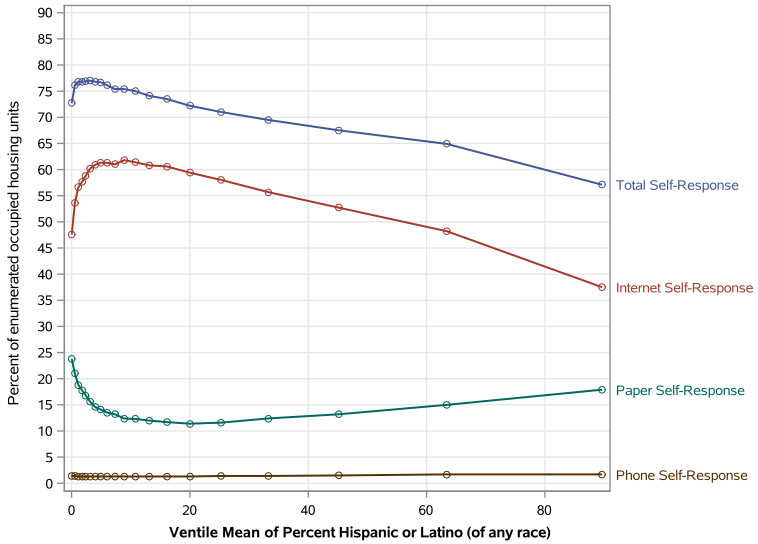


Figure 6.9 Mean Self-Response return rate by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent non-Hispanic White alone and (b) percent non-Hispanic Black or African American alone.

SOURCE: 2020 Census Self-Response Return Rate Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0224.

Copyright National Academy of Sciences. All rights reserved.



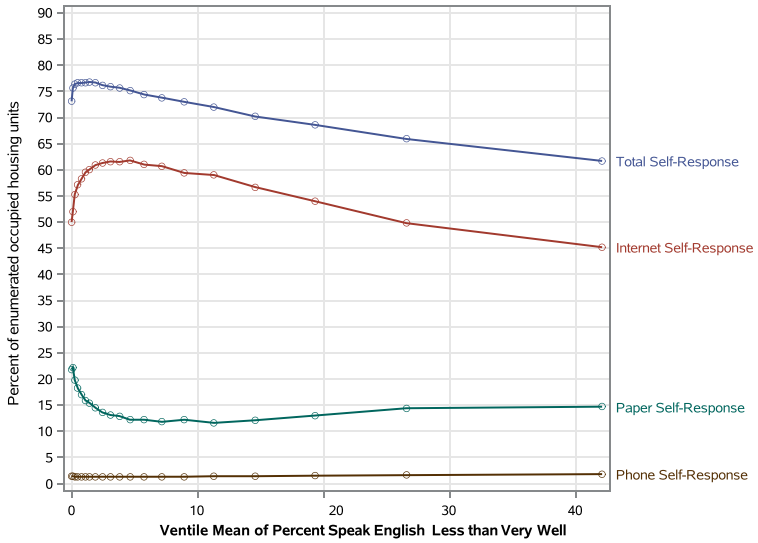
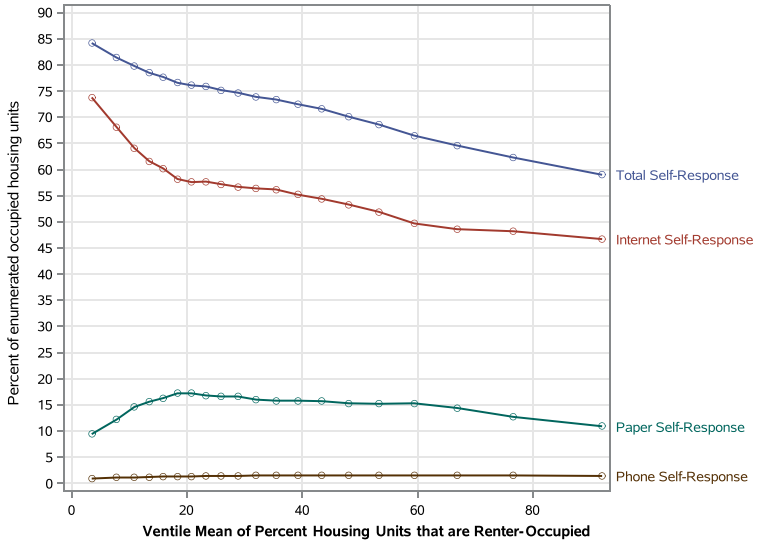


Figure 6.10 Mean Self-Response return rate by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent Hispanic or Latino (any race) and (b) percent speak English less than very well.

SOURCE: 2020 Census Self-Response Return Rate Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0224.

Copyright National Academy of Sciences. All rights reserved.



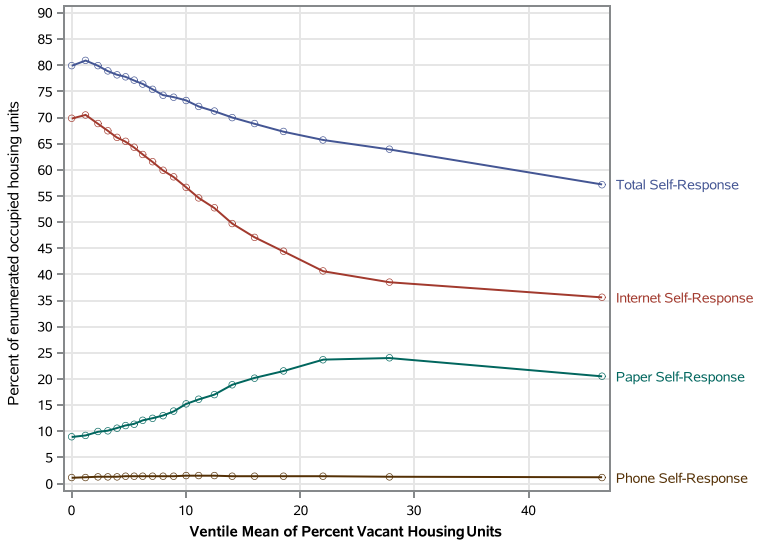


Figure 6.11 Mean Self-Response return rate by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent housing units renter occupied and (b) percent vacant housing units.

SOURCE: 2020 Census Self-Response Return Rate Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0224.

Copyright National Academy of Sciences. All rights reserved.

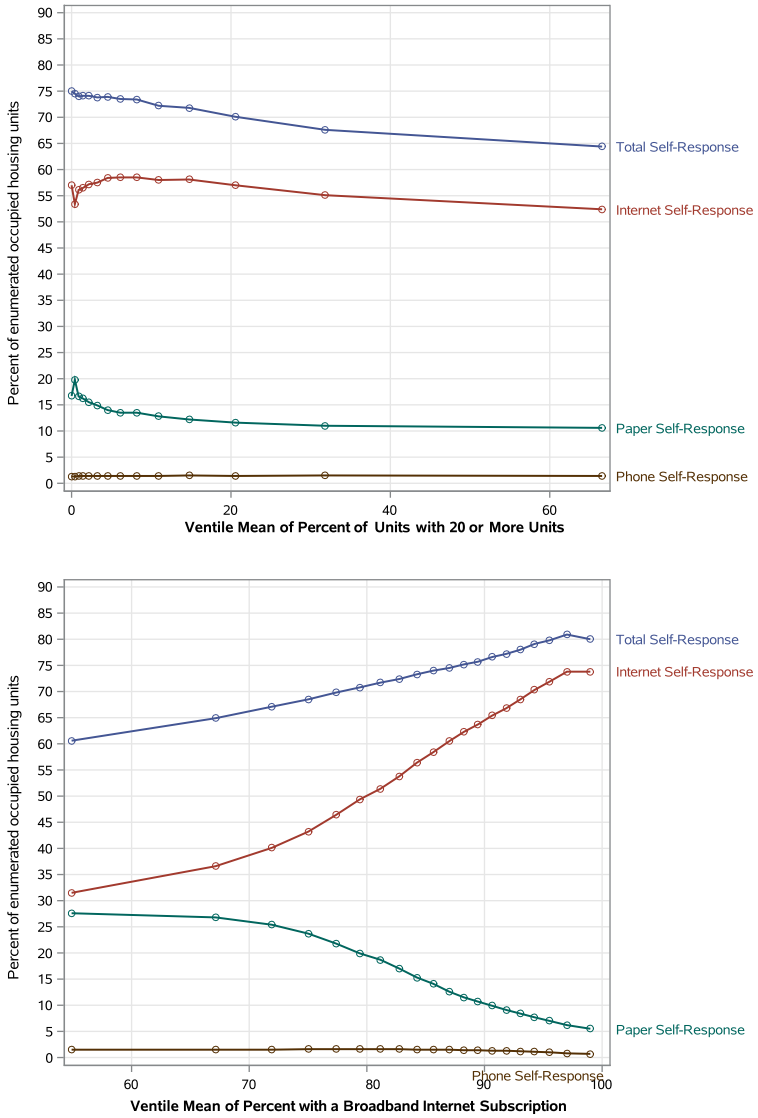


Figure 6.12 Mean Self-Response return rate by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent structures with 20 or more units and (b) percent with broadband internet subscription.

SOURCE: 2020 Census Self-Response Return Rate Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0224.

Copyright National Academy of Sciences. All rights reserved.

6.3 CONCLUSIONS AND RECOMMENDATIONS

The majority of self-response questionnaires in the 2020 Census were submitted via the internet; paper remained an important response channel, and a small share of returns by telephone complemented the results. Stated as such, this result is so obvious and unsurprising that it is easy to forget that adoption of the internet as the primary response channel was a seismic change for U.S. census taking, that internet response was eschewed entirely in the 2010 Census, and that it took time and concerted effort for the Census Bureau to put a firm stake in the ground for internet data collection in the decennial census. The U.S. 2020 Census experience built from international experience with internet collection in censuses, in particular Canada's refinement of "wave methodology" for planning its sequence of mailings and other communications to maximize response yield in the quinquennial census. Using its own survey information and modeling, the Census Bureau divided its Self-Response TEAs into Internet First and Internet Choice segments, with still-nonresponding households in the latter being slated to receive a paper questionnaire earlier than those in the former. The 2020 Census mailing strategy was affected by the COVID-19 pandemic adjustments, including last-minute addition of blanket reminder mailings, and the internet response channel weathered the risky, last-second reversion to the home-grown Primus system for data collection. However, it is difficult to imagine that the 2020 Census could have been completed without internet data collection as a foundation, including the provision for Non-ID response without reference to a physical-mail access code.

> *Conclusion 6.1: Going from a paper-only 2010 Census to a majority Internet Self-Response 2020 Census is a commendable achievement, and it is particularly commendable that the U.S. Census Bureau's internally developed Primus system was able to be swapped in immediately prior to the census with excellent performance throughout the count. (Primus was the backup system pressed into primary service after concerns surfaced regarding the outsource-developed system.) The distinction between Internet First and Internet Choice contact strategies was a reasonable starting position.*

> *Conclusion 6.2: With available return rate and operational data, it was impossible to fully evaluate the 2020 Census mailing and contact strategy beyond gross effects, such as somewhat earlier elevated levels of paper-questionnaire takeup in Internet Choice areas, and it was also impossible to disentangle effects of messaging, delivery, and response mode choices brought about by operations during the COVID-19 pandemic.*

Proportional distribution of enumeration modes varies geographically and with social, economic, and housing characteristics. Given that some modes

Copyright National Academy of Sciences. All rights reserved.

of enumeration result in poorer-quality data, if a larger proportion of a population is enumerated by modalities associated with poorer data quality, it is implied that the data quality for those populations will be poorer. While the Census Bureau systematically imputes missing values, the error associated with imputation implies that data quality is likely to be better for those items collected directly from householders.

> *Conclusion 6.3: Mode configurations varied substantially across geography, and this variation was associated with social, economic, and housing characteristics. People in areas with higher concentrations of minority populations, lower socioeconomic status, and less-stable housing were more likely to be enumerated in Nonresponse Followup and by modalities with high item nonresponse. The quality of 2020 Census data (as indicated by higher item nonresponse) is poorer in areas with higher concentrations of minorities, lower socioeconomic status, and less stable housing. This raises concerns about data equity, with progressively less accurate and less complete data being collected differentially by race and ethnicity, socioeconomic status, and housing tenure.*

Another issue with less-than-accurate data on age, race, and ethnicity is that, for the balance of the decade, census data become the statistical foundation for many programs. In addition to this decennial data foundation varying in accuracy across age, race, ethnicity, and probably sex, introducing differential privacy compounds questions about the accuracy and utility of 2020 Census data for some programs and small-area analyses. Programs that use census data and the demographic-detail data as their data foundation will be less than accurate, and the amount of error present in census data will likely be compounded for programs such as postcensus population estimates and projections. Census demographic-detail data are used by numerous statistical programs as denominators for calculating age, sex, race, and ethnicity-specific rates. Less-than-accurate denominators can potentially minimize or amplify true differences between geographies and subpopulations—an issue that is even greater for small geographic areas or analyses of small subpopulations.

> *Conclusion 6.4: The magnitude of uncertainty in terms of the accuracy of counts of population characteristics as a function of item nonresponse is unknown. The variation in enumeration modality associated with social, economic, and housing characteristics results in differences in data quality by these characteristics. We do not understand the overall impact of item nonresponse and the associated imputation of values on the accuracy of data for the total population or subpopulations. A better understanding of the differential impact of imputation*

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

across subpopulations would potentially enable prioritized targeting of groups, to improve self-reporting.

Recommendation 6.1: The U.S. Census Bureau should continue to investigate the performance of the contact strategies used in the 2020 Census, to increase overall self-response (particularly via the internet) in 2030, more in concordance with international experience with online census response. This investigation should include review of the effectiveness of the 2020 Census's segmentation into Internet First and Internet Choice strategies, as well as the timing and the operative contact/enumeration strategy (e.g., Internet First or Internet Choice mailings or Update Leave delivery) underlying successful Non-ID enumerations in 2020.

Recommendation 6.2: The U.S. Census Bureau should engage in further research to more specifically identify social, economic, and housing characteristics associated with less-than-accurate forms of enumeration, and should also research related communication and operational strategies that would improve self-response and associated data accuracy for those population and housing segments associated with poorer-quality census data in 2020.

Recommendation 6.3: The U.S. Census Bureau should assess the magnitude of the impact of imputation on the accuracy of 2020 Census data. Assessment of the differential impact of imputation on subpopulations should be used to prioritize research into strategies to improve self-response.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

– 7 –

Nonresponse Followup

In this chapter, we explore aspects of the 2020 Census Nonresponse Followup (NRFU) related to the potential impact on aspects of census quality. We begin by examining the operational phases of NRFU and how several modes of enumeration contributed to the completion of the NRFU phase of the 2020 Census operation. We then examine geographic variation in the composition of NRFU modes of enumeration, followed by an exploration of associations between NRFU enumeration modalities and demographic, economic, and housing characteristics.

7.1 OVERVIEW OF 2020 NONRESPONSE FOLLOWUP OPERATION AND CHALLENGES

The U.S. Census Bureau identifies types of enumeration areas (TEAs) to guide enumeration strategies. Two types of TEAs are relevant for NRFU; Self-Response and Update Leave (UL). Self-Response is designed to occur in areas where the majority of housing units have mail delivered to the physical location of the housing unit. Self-Response was the primary enumeration methodology for the 2020 Census. UL is designed to occur in areas where the majority of housing units either do not have mail delivered to the physical location of the housing unit, or the mail-delivery information for the housing unit cannot be verified.

Through the 2020 NRFU operation, census takers conducted in-person interviews to enumerate households that had not yet responded online, by mail, or by phone. Sixty-seven percent of households self-responded, and 33% were

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

enumerated through NRFU. Due to COVID-19, the start of NRFU was shifted by approximately three months to mid-July. With this later start, the Early NRFU operation (Phase 1), designed to reach college students before the end of the spring semester, was no longer conducted. Given this change, the NRFU operation began with a soft launch on July 16 in specific areas of the country. By August 9, NRFU was being conducted nationwide, and it continued for nearly two more months until October 15 when both production operations and quality control concluded.[1]

Original plans for the phased approach to NRFU set both case-completion and fixed-date criteria, by which census field supervisor areas (CFSAs) could transition between phases. On the case-completion front, early work suggested that CFSAs were reaching the 60% and 85% thresholds to switch to Phase 2 or Phase 3, respectively, earlier than expected, with "many cases" not yet having a first contact attempt. Accordingly, the threshold was reset from 60% to 85% complete to transition to Phase 2, and from 85% to 90% to transition to Phase 3. The original dates for eligibility to shift phases were made inoperable by the delay in starting NRFU and the phased restart of operations across Area Census Office jurisdictions. Ultimately, September 4, 2020, became the new date of eligibility for Phase 2 and September 11, 2020, for Phase 3 (Fontenot, 2021d).

In early September, the Census Bureau was anticipating direction from the U.S. Congress to complete the 2020 Census apportionment count by the legal deadline of December 31. The operation implemented a series of changes to reduce the NRFU timeline to meet this deadline. Shortly after implementing the changes, a federal judge issued a court order that reversed those changes. One of the changes was the number of contact attempts made for NRFU Reinterview (NRFU RI) and Self-Response Quality Assurance (SRQA) cases. The number of contact attempts was reduced from six to three; this situation existed for approximately two weeks before a court order set the criteria back to six. During that period, some NRFU RI and SRQA cases were closed after three attempts and processed as such. The majority of NRFU RI and SRQA cases did receive six attempts. Phase 2 and Phase 3 of the NRFU contact strategy were largely predicated on mid-decade research findings indicating that high-performing enumerators significantly contributed to the completion of difficult cases. However, during the 2020 NRFU operation, federal court orders blocked the Census Bureau from reducing field staff, including low-performing staff. Completion date changes and the restriction on dismissing low-performing staff prevented the new phased approach to NRFU from being conducted exactly as planned.

[1]See https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/operational-adjustments.html.

Copyright National Academy of Sciences. All rights reserved.

The 2020 Census NRFU operation was also very different from its predecessors in terms of its use of administrative records data, not only to flag potential vacant and delete-worthy cases early on (for removal from fieldwork after one confirmatory field visit) but as a means for enumerating households if information of sufficient quality existed for a nonresponding household.

As noted, during late stages of NRFU, there was a period of uncertainty about when data collection would end. The unanticipated need to adjust NRFU operations in response to COVID-19 led the Census Bureau to seek to complete data collection at the end of October 2020. Congress did not approve this, and, in early August, September 30 became the target end date for data collection. An injunction forcing the end date back to October 31 was enacted but shortly thereafter escalated to the U.S. Supreme Court, which stayed the injunction in an unsigned order, paving the way for census field operations and data collection to end on October 15. This uncertainty of ending data collection for one of the most sizable and significant logistical operations in the United States forced the Census Bureau to be nimble and adjust staffing and operations numerous times. It also raises questions about the potential impact on the quality of the enumerations during closeout.

Issues of data quality noted in Chapter 6 are also significant for NRFU data. There are numerous challenges in collecting quality census data from households. Item nonresponse is a significant challenge that often results in inaccurate values being imputed (see Chapter 3). Given that item nonresponse and associated data quality vary from one mode of enumeration to the next, we examined variation in the mode of enumeration across NRFU phases. We conducted a descriptive analysis of NRFU modalities for the nation, counties, and census tracts. We also examined the clustering of NRFU modality configurations across counties. Finally, we examined geographic variation in NRFU response mode and the association with demographic, social, and housing characteristics to assess potential data-quality differences across these characteristics.

7.2 DESCRIPTIVE ANALYSIS OF NONRESPONSE FOLLOWUP ENUMERATION BY MODALITY AND PHASE

The Census Bureau provided the panel with data on NRFU by phase for the nation, counties, and census tracts. The data included information about the number of households enumerated by NRFU enumeration modalities for each level of geography. These data were used to assess modal patterns of enumeration and associations of these patterns with demographic, economic, and housing characteristics.

Table 7.1 summarizes enumeration modes for each phase, at the national level. Of the national total NRFU enumerations, 42% were by household

Copyright National Academy of Sciences. All rights reserved.

Table 7.1 National Summary of 2020 Census Nonresponse Followup Resolutions by Major Phase

	Phase 1		Phase 2		Phase 3		
	N	Percent	*N*	Percent	*N*	Percent	Total
Starting Workload	59,280,000		11,430,000		6,479,000		
– Self-Response Removals	5,732,000		202,000		184,000		6,119,000
Interview/Observation	19,150,000	45.5%	1,606,000	33.8%	1,666,000	27.9%	22,430,000
Occupied	12,110,000		1,233,000		1,327,000		
Vacant	1,566,000		92,000		90,000		
Delete	5,481,000		281,000		248,000		
Proxy	14,550,000	34.5%	2,536,000	53.4%	3,413,000	57.1%	20,500,000
Occupied	3,944,000		1,314,000		1,853,000		
Vacant	7,570,000		925,000		1,116,000		
Delete	3,038,000		297,000		445,000		
Administrative Records	8,213,000	19.5%	290,000	6.1%	506,000	8.5%	9,009,000
Occupied	6,251,000		173,000		416,000		
Vacant	1,629,000		101,000		78,500		
Delete	333,000		15,000		11,000		
Special	200,000	0.5%	315,000	6.6%	394,000	6.6%	908,000
Population Count Only	162,000		309,000		376,000		
Special Closeout	38,000		5,200		17,000		
NRFU Resolutions in Phase	42,120,000		4,746,000		5,978,000		52,840,000
Left Unresolved	11,430,000		6,479,000		317,000		317,000

NOTES: NRFU, Nonresponse Followup.

SOURCE: Tabulation from Nonresponse Followup Resolutions by Phase dataset, 2020 Census operational data. See Disclosure Review Statement; CBDRB-FY23-0214.

Copyright National Academy of Sciences. All rights reserved.

interview, 39% were by proxy, and 17% were by administrative records, with only 2% enumerated as special (count only and special closeout). This distribution varies across phases, with most NRFU enumerations in Phase 1 being household interviews (45.5%) followed by proxy (34.5%), administrative records (19.5%), and special accounting for only 0.5%. For Phases 2 and 3, proxy is the more common mode (53.4% and 57.1%, respectively), followed by household interviews (33.8% and 27.9%), and administrative records (6.1% and 8.5%). Thus there was a smaller proportion of administrative record enumerations after Phase 1 and a more significant proportion of enumerations that were by proxy and special. Differences in the mode distribution for Phases 2 and 3 were minor, indicating that, at the end of NRFU, there were not any dramatic shifts toward lower-quality modes of enumeration. The graphic depiction of the distribution of enumeration mode by phase at the CFSA unit of analysis (Figure 7.1) is consistent with the distribution observed for national data.

To further examine differences and shifts in enumeration mode by phases, we combined Phases 2 and 3, constructed the percent of enumerations by mode, and then subtracted the percent of enumerations by mode in Phase 1 for counties (Figure 7.2). Thus, positive values for a mode indicate a shift toward that mode in the later phases, and negative values indicate a change away from that mode. We observed a shift away from household interviews and administrative records during Phases 2 and 3 and toward enumeration by proxy. This pattern is further illustrated in Figures 7.3 (county) and 7.4 (census tract), in which Phase 1 and combined Phases 2 and 3 are compared by the percentage of enumerations performed using each mode.

7.3 CLUSTER ANALYSIS OF VARIATION IN NONRESPONSE FOLLOWUP RESOLUTIONS BY PHASE

To better understand commonalities in the distributions of NRFU enumeration modalities across geography, we generated three clusters using k-means clustering for NRFU enumeration modality configurations across counties and census tracts. K-means clustering is a vector quantization method, originally from signal processing, that aims to partition n observations into k clusters, in which each observation belongs to the cluster with the nearest mean (cluster centers or cluster centroid), serving as a prototype of the cluster.

Results (Figure 7.5) suggest that Cluster 1 seems to have high administrative records takeup in Phase 1 and a significant gap in percent occupied by proxy between Phases 1 and 2–3. Cluster 2 has generally higher vacant by proxy and lower occupied by interview in both phase groups. Finally, Cluster 3 seems to have higher levels of deletes. We then mapped the three clusters by county (Figure 7.6) and census tracts (Figure 7.7). The county- and census tract-

Copyright National Academy of Sciences. All rights reserved.

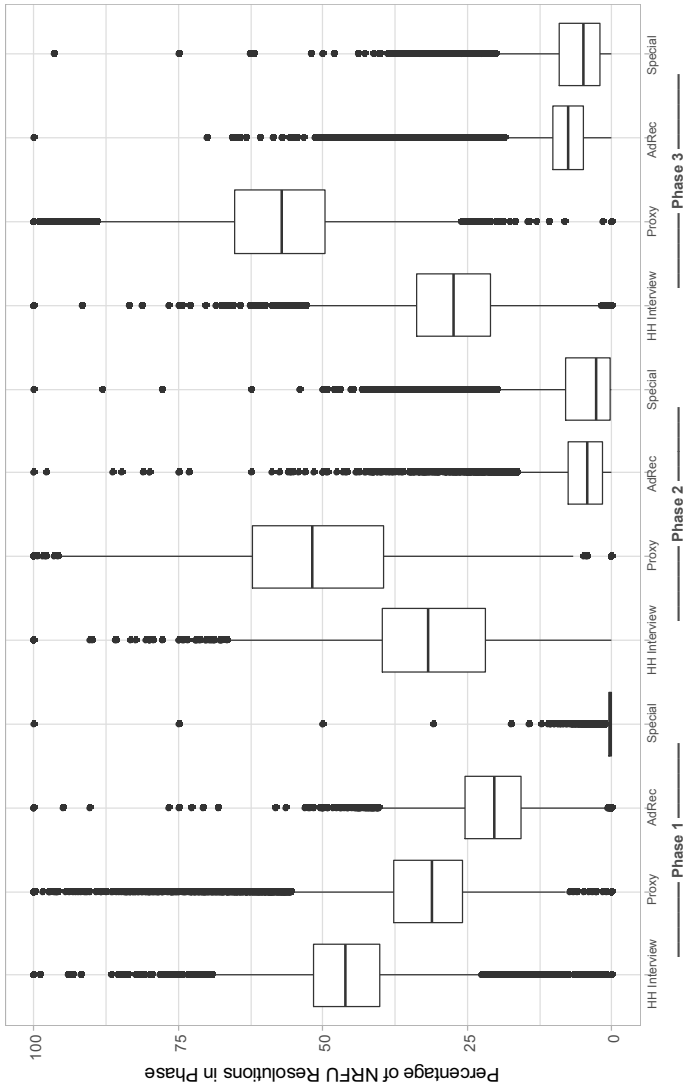


Figure 7.1 Nonresponse Followup resolutions by household interview/observation, proxy, administrative records, and special means, by operational phase, 2020 Census Field Supervisor Areas.

NOTES: AdRec, administrative records; HH, householder; NRFU, Nonresponse Followup.
SOURCE: NRFU Resolutions by Phase Dataset. See Disclosure Review Statement; CBDRB-FY23-0171.

Copyright National Academy of Sciences. All rights reserved.

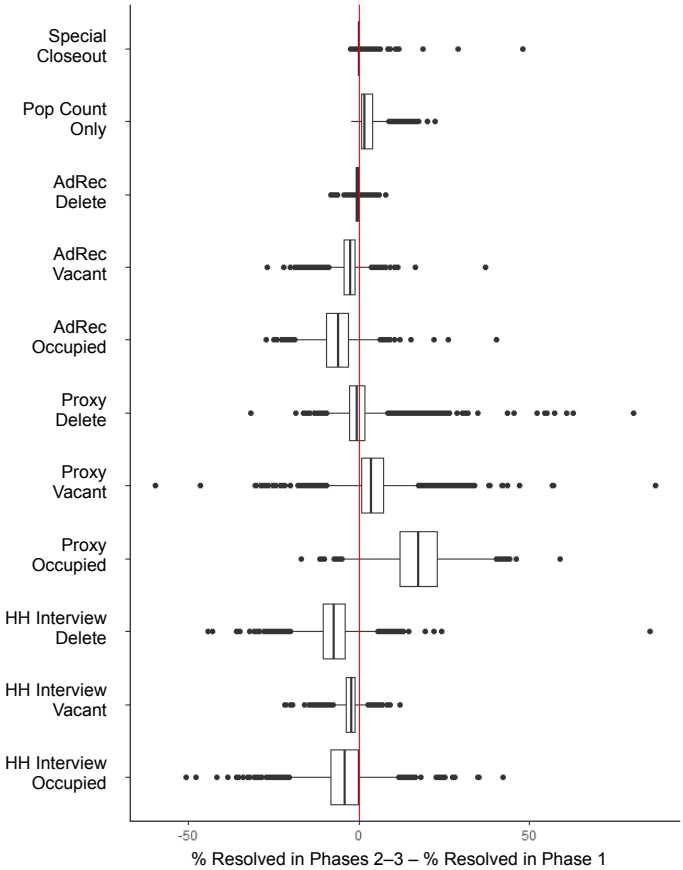


Figure 7.2 Percent of workload resolved in Phases 2–3 minus percent resolved in Phase 1, by major resolution mode and outcome, by county.

NOTES: AdRec, administrative records; HH, householder; NRFU, Nonresponse Followup.

SOURCE: NRFU Resolutions by Phase Dataset. See Disclosure Review Statement; CBDRB-FY23-0171.

Copyright National Academy of Sciences. All rights reserved.

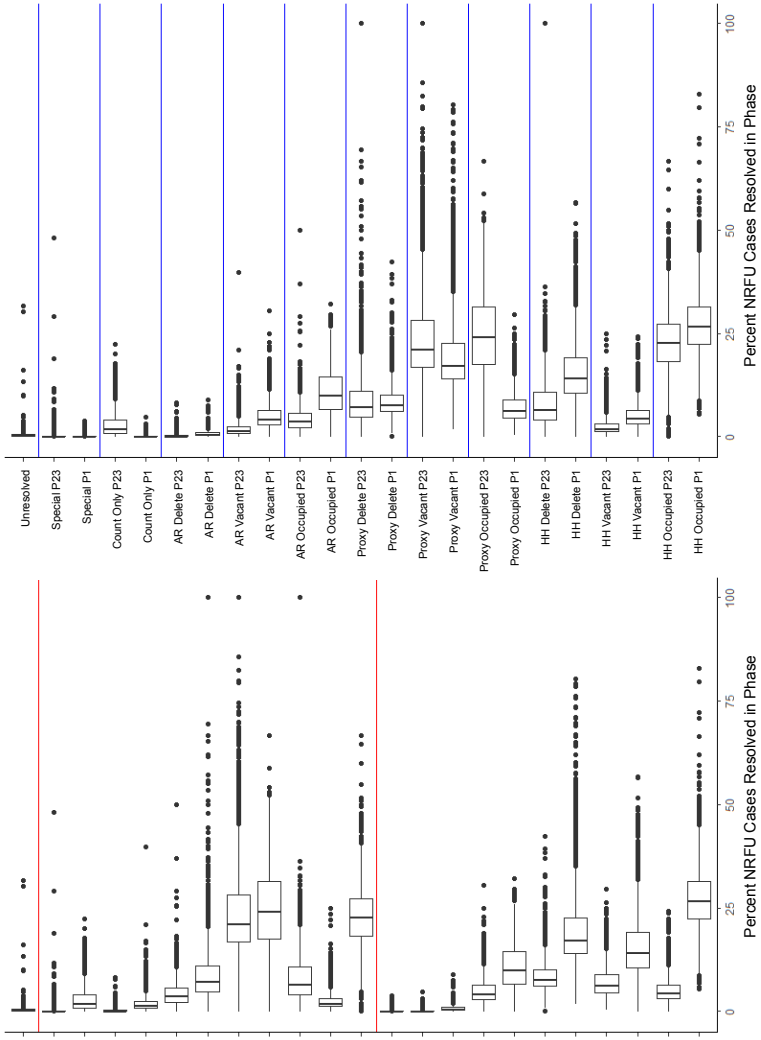


Figure 7.3 Summary Nonresponse Followup resolutions by phase (combining Phases 2 and 3), mode, and outcome, by county.

NOTES: AR, administrative records; HH, householder; NRFU, Nonresponse Followup. Left and right panels differ in order of stacking the boxes, separate by phase (1 and 2–3 combined) at left and by mode/outcome at right.
SOURCE: NRFU Resolutions by Phase Dataset. See Disclosure Review Statement; CBDRB-FY23-0171.

Copyright National Academy of Sciences. All rights reserved.

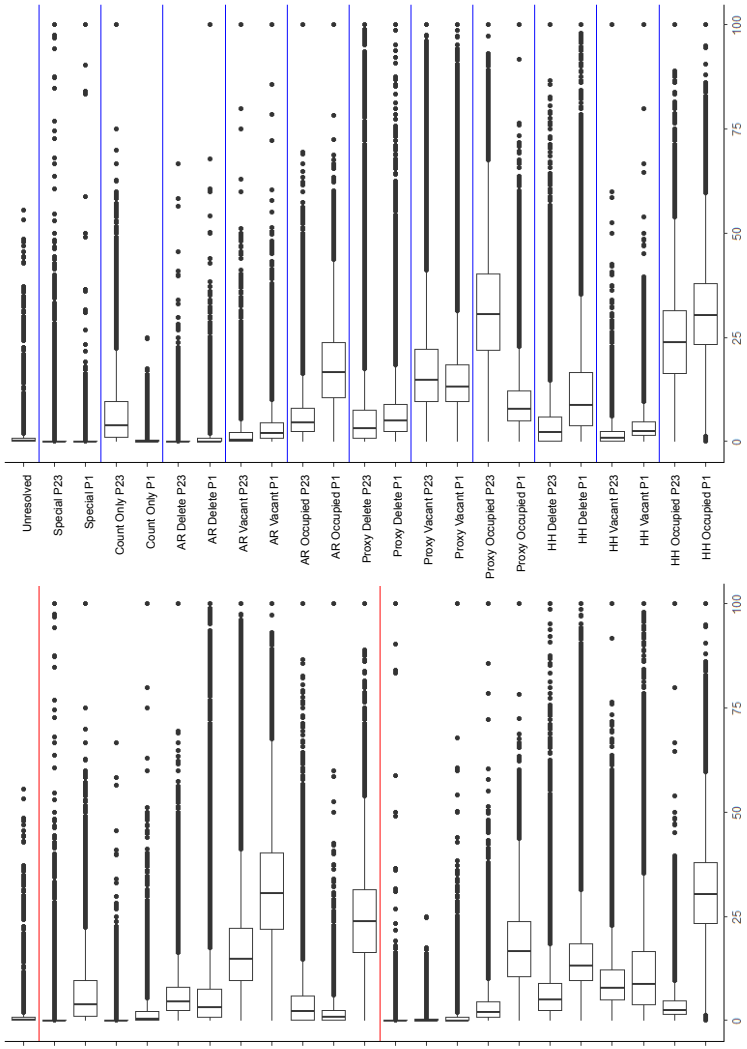


Figure 7.4 Summary Nonresponse Followup resolutions by phase (combining Phases 2 and 3), mode, and outcome, by tract.

NOTES: AR, administrative records; HH, householder; NRFU, Nonresponse Followup. Left and right panels differ in order of stacking the boxes, separate by phase (1 and 2–3 combined) at left and by mode/outcome at right.

SOURCE: NRFU Resolutions by Phase Dataset. See Disclosure Review Statement; CBDRB-FY23-0171.

Copyright National Academy of Sciences. All rights reserved.

level maps indicate that Cluster 1 tends to be in more urbanized areas whereas Clusters 2 and 3 appear to be in more rural areas.

7.4 NONRESPONSE FOLLOWUP ENUMERATION MODALITIES AND DEMOGRAPHIC, ECONOMIC, AND HOUSING CHARACTERISTICS

As with self-response, areas with a higher proportion of households enumerated with modes that result in low item nonresponse will have higher quality data than those with a higher proportion of enumerations by modes with high item nonresponse. Geographic variation in the quality of collected data is likely associated with variations in populations' social, economic, and housing characteristics as response modality varies along these characteristics.

We examined the association between 2020 Census NRFU enumeration modalities and social, economic, and housing characteristics for census tracts in the United States. Two research questions lie at the core of this effort. First, what is the nature of the associations between the proportion of housing units enumerated by the various NRFU modalities in the 2020 Census and census tracts' social, economic, and housing characteristics? Second, what do associations between 2020 Census census-tract NRFU modalities and social, economic, and housing characteristics indicate about variation in the quality of 2020 Census data across social, economic, and housing characteristics of populations?

For this analysis, we used the same NRFU Resolutions by Phase dataset but opted not to distinguish between the phases of the NRFU operation. To make our analysis comparable with the work in Chapter 6, we further restricted our attention to NRFU resolutions of units as "occupied" (not vacant or delete) for the principal NRFU modalities of household interview, administrative records, and proxy. We derived tract-level percentages for these three modalities (as well as population count only, special closeout, and unresolved); scaled those percentages to approximate percentage of enumerated occupied housing units by multiplying by (100 − Self-Response return rate); and merged these data with the same public 2016–2020 5-year sample American Community Survey (ACS) variables as used in Chapter 6. From the resulting data, we calculated zero-order bivariate correlation coefficients between each variable in the merged dataset. The correlation coefficients between each modality and each ACS variable are graphed from high to low for each modality.

To further examine associations between NRFU response modalities and demographic, economic and housing characteristics from the ACS, we ranked census tracts for each variable. Then we grouped ranked tracts in up to 20 categories (ventiles), from highest to lowest. We dropped census tracts with missing data for an ACS variable. Means of each ACS variable were calculated

Copyright National Academy of Sciences. All rights reserved.

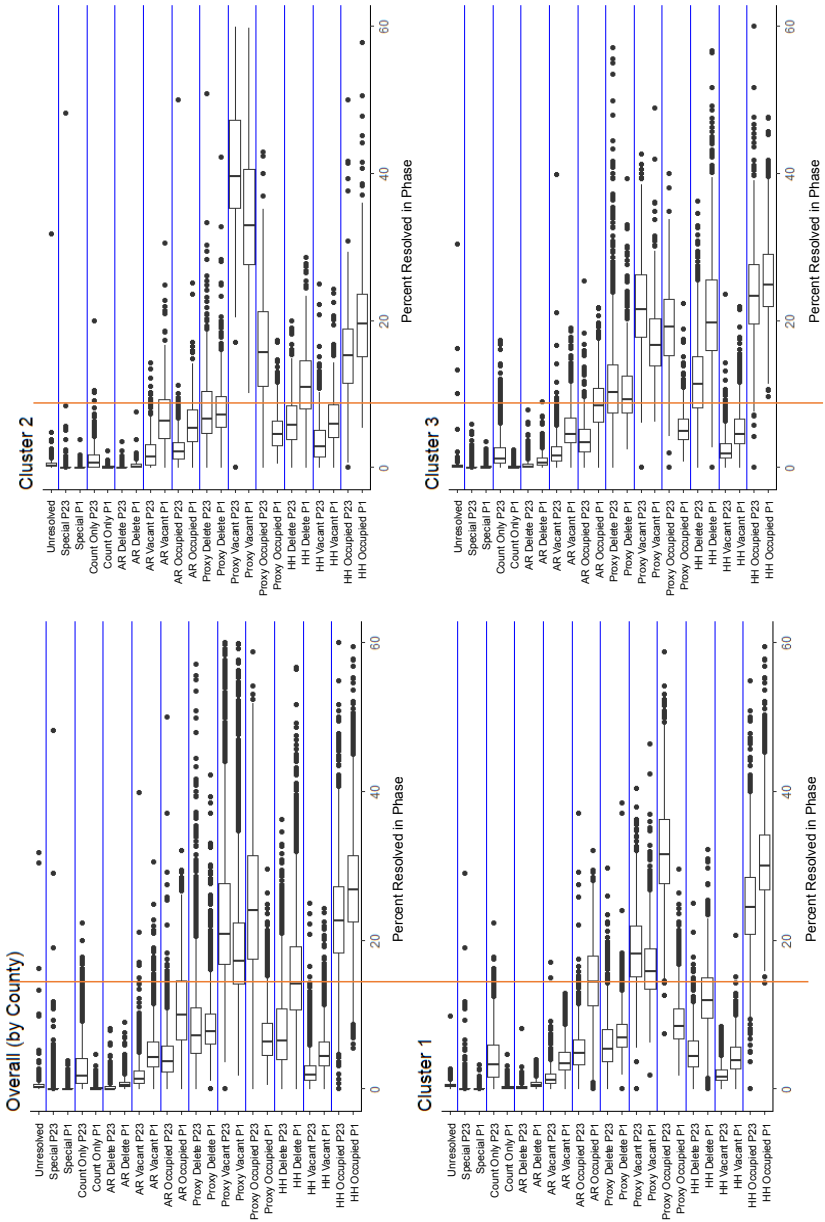


Figure 7.5 K-means clustering of county-level Nonresponse Followup resolution composition percentages into three clusters.

NOTES: AR, administrative records; HH, householder; NRFU, Nonresponse Followup.

SOURCE: NRFU Resolutions by Phase Dataset. See Disclosure Review Statement; CBDRB-FY23-0171.

Copyright National Academy of Sciences. All rights reserved.

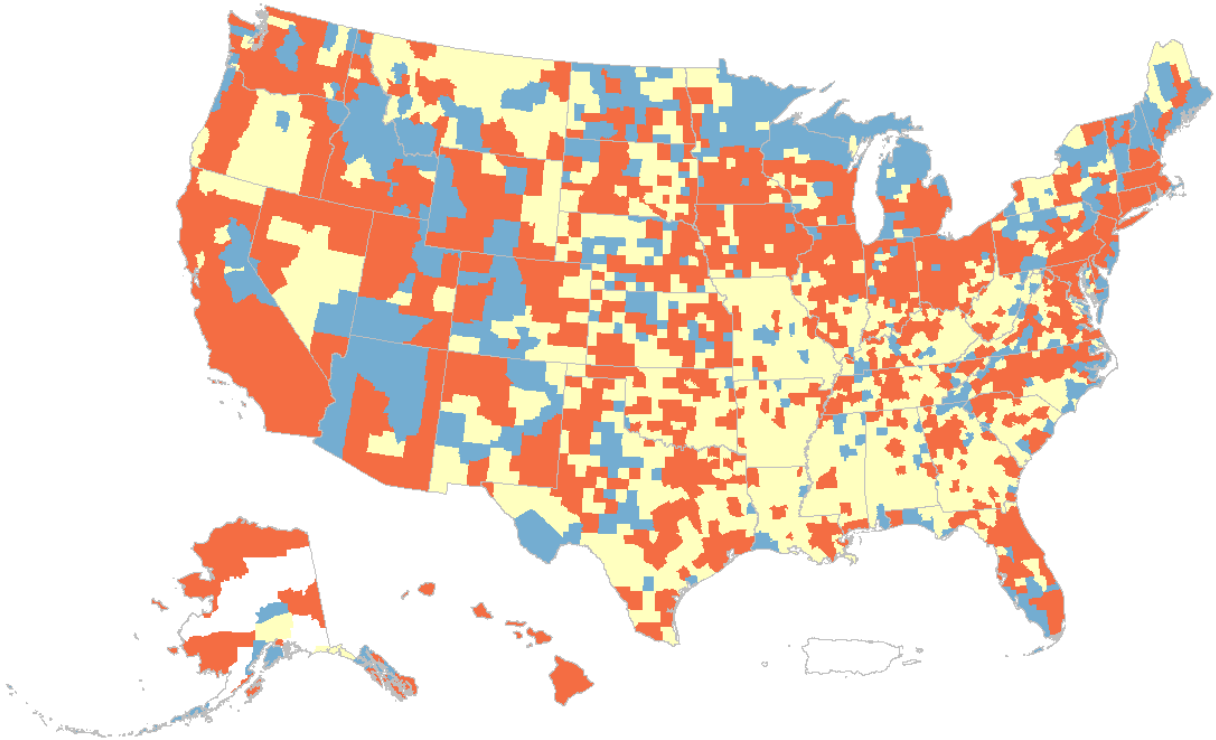


Figure 7.6 Geographic distribution of *k*-means clustering of Nonresponse Followup resolution composition percentages, by county.

NOTE: Cluster 1 (red) is generally characterized by high administrative records takeup in Phase 1 and a large gap in percent occupied by proxy between Phases 1 and 2–3; cluster 2 (yellow) is characterized by generally higher vacant by proxy and lower occupied by interview in both phase groups; cluster 3 (blue) is characterized by generally higher levels of deletes.

SOURCE: NRFU Resolutions by Phase Dataset. See Disclosure Review Statement; CBDRB-FY23-0171.

Copyright National Academy of Sciences. All rights reserved.

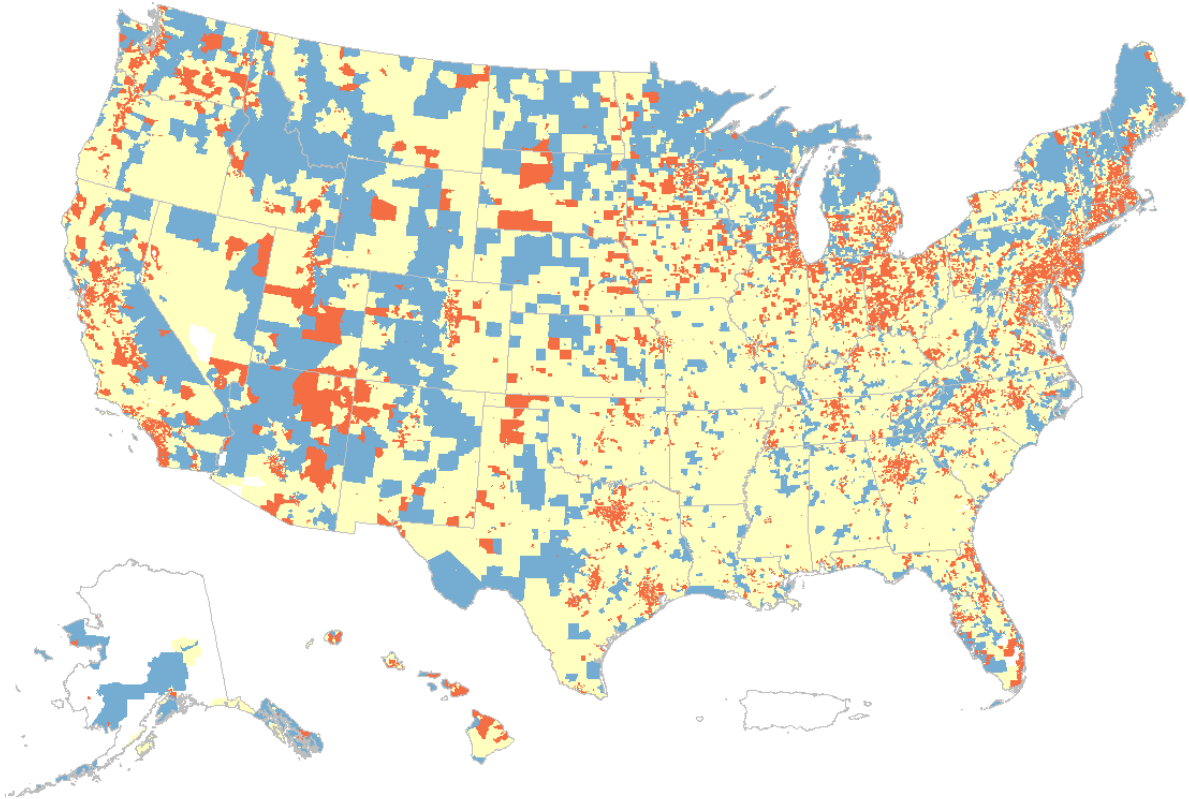


Figure 7.7 Geographic distribution of *k*-means clustering of Nonresponse Followup resolution composition percentages, by tract.

NOTE: Cluster 1 (red) is generally characterized by high administrative records takeup in Phase 1 and a large gap in percent occupied by proxy between Phases 1 and 2–3; cluster 2 (yellow) is characterized by generally higher vacant by proxy and lower occupied by interview in both phase groups; cluster 3 (blue) is characterized by generally higher levels of deletes.

SOURCE: NRFU Resolutions by Phase Dataset. See Disclosure Review Statement; CBDRB-FY23-0171.

Copyright National Academy of Sciences. All rights reserved.

for each ventile category, as was the mean of the percent of households enumerated using each NRFU enumeration mode. Line graphs were generated from the resulting data, illustrating the variation in modality from one ventile to the next and across all ACS ventiles.

Our data-quality assessment for NRFU is focused on household interviews, administrative records, proxy interviews, and count only. As described in Chapter 3, household interviews have fewer item nonresponses for race and Hispanic origin than do other NRFU modalities. Still, administrative records have fewer item nonresponses for age. Thus, for NRFU, household interviews provide the most accurate responses, followed by administrative records, and proxy interviews appear to have the highest percentage of item nonresponse.

7.4.1 Nonresponse Followup Correlations with Social and Economic Characteristics of Census Tracts

The correlation between ACS variables and the percentage of enumeration from household interviews is somewhat similar to the pattern observed with Paper Self-Response, as noted in Chapter 6 (see Figures 7.8–7.11). Lower socioeconomic status (SES) characteristics appear to be positively associated with household interviews. This is also true for other NRFU modalities. All NRFU modalities tend to be positively associated with higher percentages of low SES characteristics and negatively associated with higher SES characteristics. Given that all NRFU modalities are more likely to have higher item nonresponse than all Self-Response modalities, populations in census tracts with a higher percent of lower SES characteristics tend to be represented by poorer-quality 2020 Census data.

7.4.2 Nonresponse Followup Ventile Plots

Figures 7.12–7.17 present analysis of the approximate percentage of enumerated occupied housing units resolved by the major NRFU modalities, in relation to ventiles of the same ACS covariates used in Chapter 6. The percentage of household interviews was higher for census tracts with larger average household sizes (see Figures 7.12–7.13). Note that the highest y-axis value of these graphs is 30%, to allow an easier review of differences between ventiles. Census tracts with smaller average household sizes were more likely to be enumerated by proxy than were tracts with larger average household sizes. Consequently, tracts with smaller household sizes will likely have poorer quality data representing them in the 2020 Census. Census tracts with smaller percentages of adults with bachelor's degrees or higher, and with lower median household income, were more likely to be enumerated by household interviews. These tracts were also more likely to be enumerated by proxy, and slightly more likely to be enumerated by count only. Thus, census tracts with lower levels of educational

Copyright National Academy of Sciences. All rights reserved.

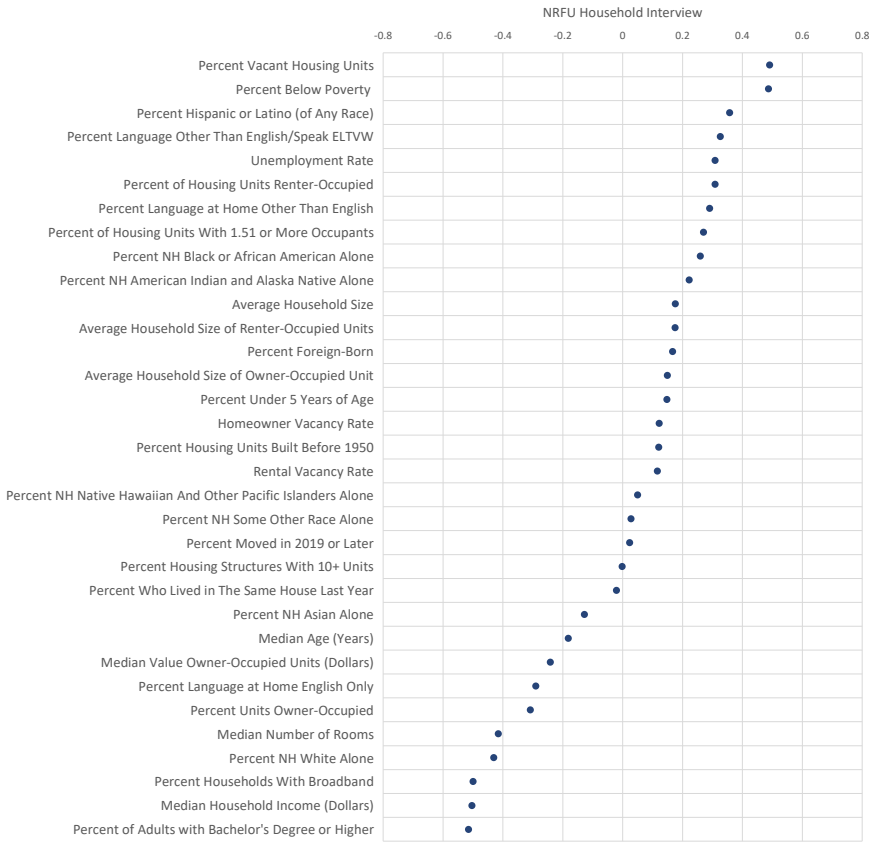


Figure 7.8 Zero-order correlations between percent Nonresponse Followup resolutions by household member interview and American Community Survey socioeconomic characteristics, census tracts.

NOTES: NH, non-Hispanic; ELTVW, English less than very well.

SOURCE: NRFU Resolutions by Phase Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0206.

Copyright National Academy of Sciences. All rights reserved.

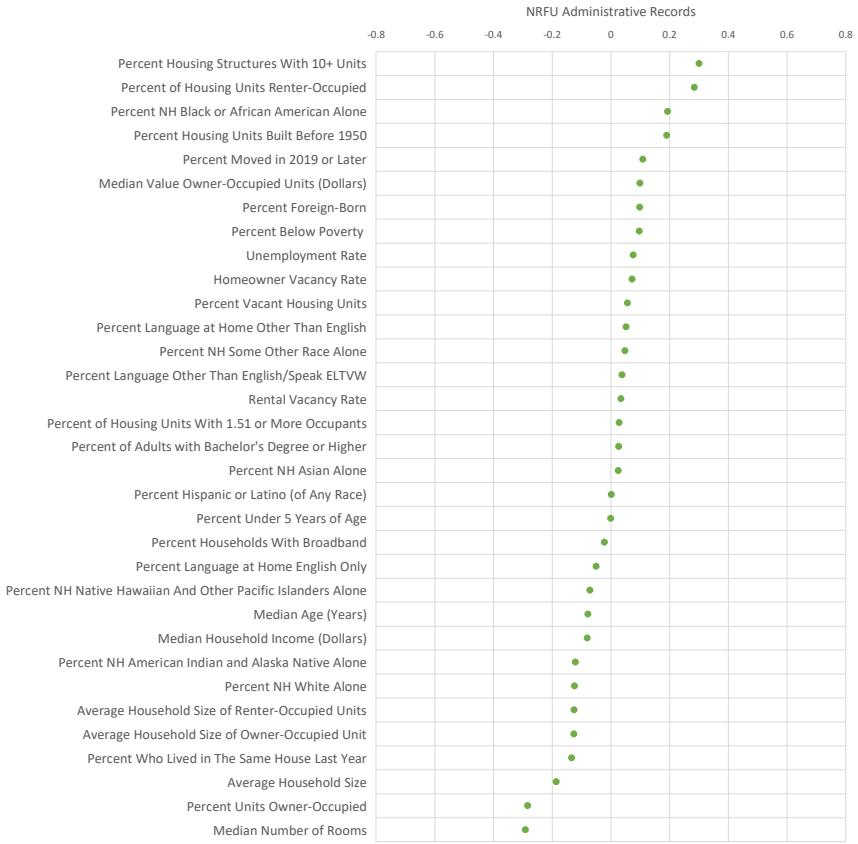


Figure 7.9 Zero-order correlations between percent Nonresponse Followup resolutions by administrative records and American Community Survey socioeconomic characteristics, census tracts.

NOTES: NH, non-Hispanic; ELTVW, English less than very well.

SOURCE: NRFU Resolutions by Phase Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0206.

Copyright National Academy of Sciences. All rights reserved.

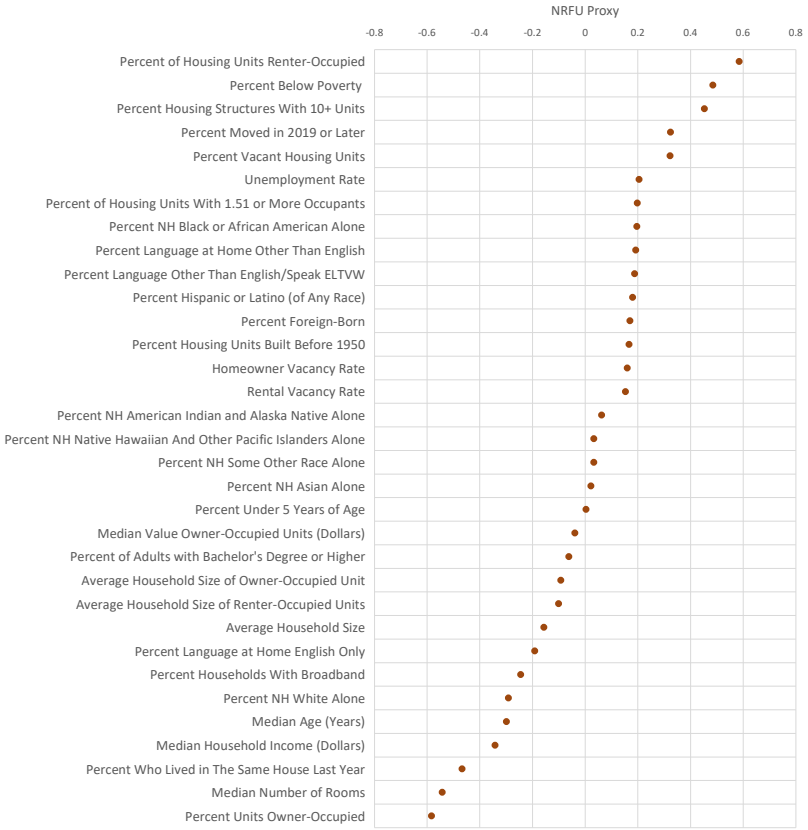


Figure 7.10 Zero-order correlations between percent Nonresponse Followup resolutions by proxy and American Community Survey socioeconomic characteristics, census tracts.

NOTES: NH, non-Hispanic; ELTVW, English less than very well.

SOURCE: NRFU Resolutions by Phase Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0206.

Copyright National Academy of Sciences. All rights reserved.

Figure 7.11 Zero-order correlations between percent Nonresponse Followup resolutions as "population count only" and American Community Survey socioeconomic characteristics, census tracts.

NOTES: NH, non-Hispanic; ELTVW, English less than very well.

SOURCE: NRFU Resolutions by Phase Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0206.

Copyright National Academy of Sciences. All rights reserved.

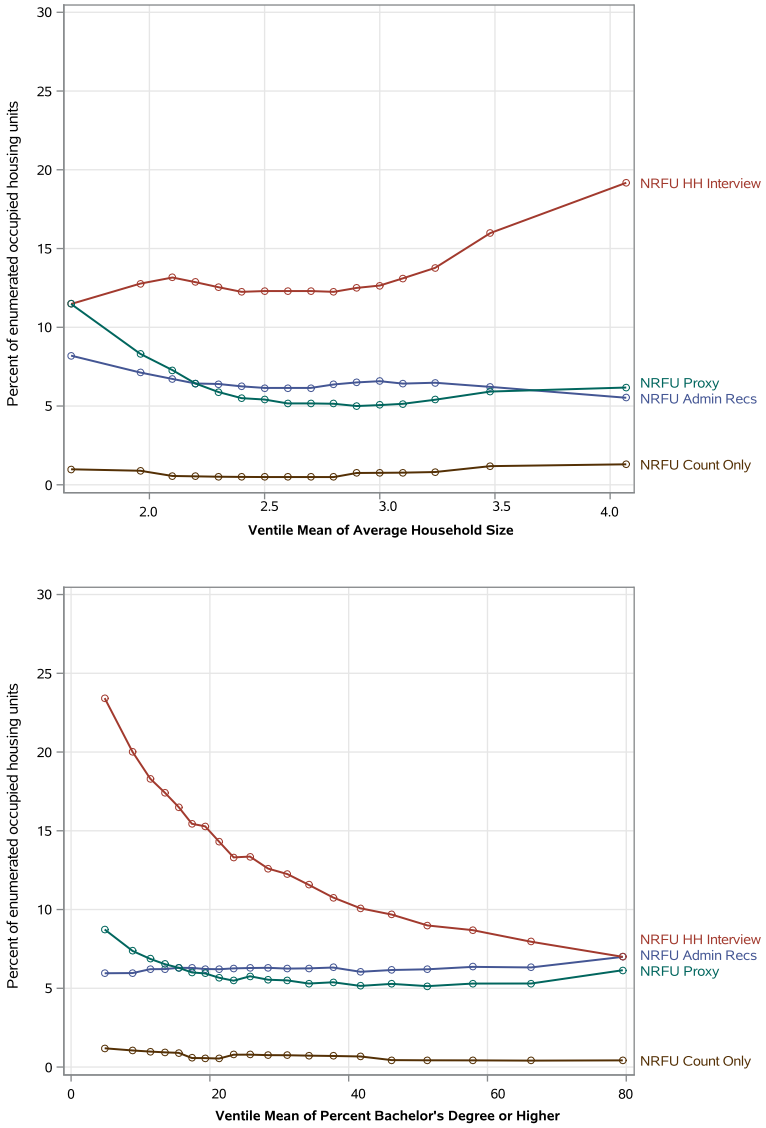


Figure 7.12 Mean percent Nonresponse Followup resolutions by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) average household size and (b) percent bachelor's degree or higher.

NOTES: NRFU, Nonresponse Followup; HH, householder.

SOURCE: NRFU Resolutions by Phase Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0224.

Copyright National Academy of Sciences. All rights reserved.

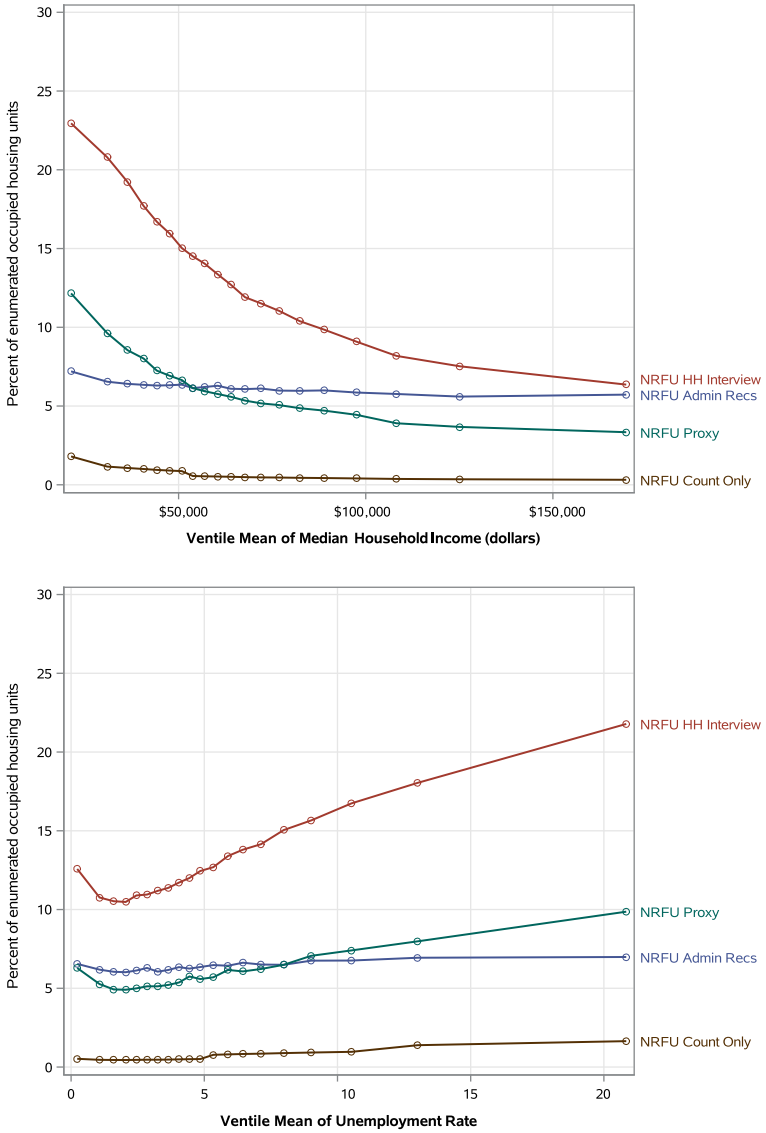


Figure 7.13 Mean percent Nonresponse Followup resolutions by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) median household income and (b) unemployment rate.

NOTES: NRFU, Nonresponse Followup; HH, householder.

SOURCE: NRFU Resolutions by Phase Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0224.

Copyright National Academy of Sciences. All rights reserved.

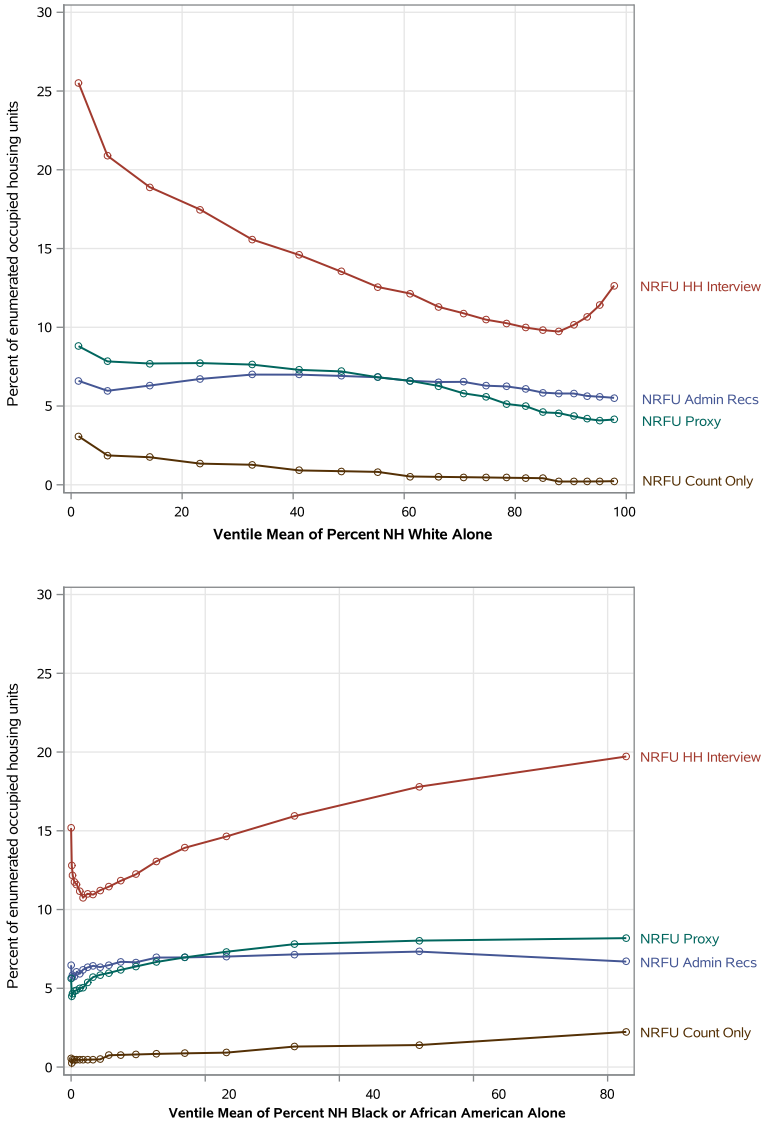


Figure 7.14 Mean percent Nonresponse Followup resolutions by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent non-Hispanic White alone and (b) percent non-Hispanic Black or African American alone.

NOTES: NRFU, Nonresponse Followup; HH, householder.

SOURCE: NRFU Resolutions by Phase Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0224.

Copyright National Academy of Sciences. All rights reserved.

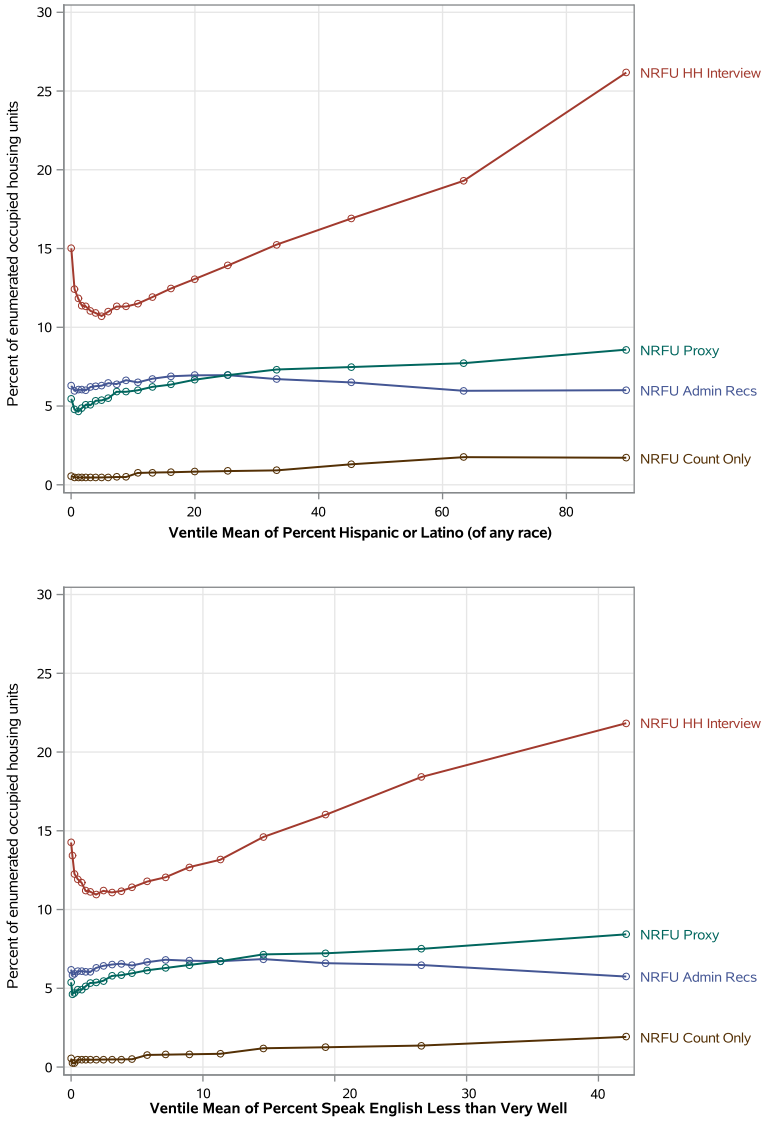


Figure 7.15 Mean percent Nonresponse Followup resolutions by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent Hispanic or Latino (any race) and (b) percent speak English less than very well.

NOTES: NRFU, Nonresponse Followup; HH, householder.

SOURCE: NRFU Resolutions by Phase Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0224.

Copyright National Academy of Sciences. All rights reserved.

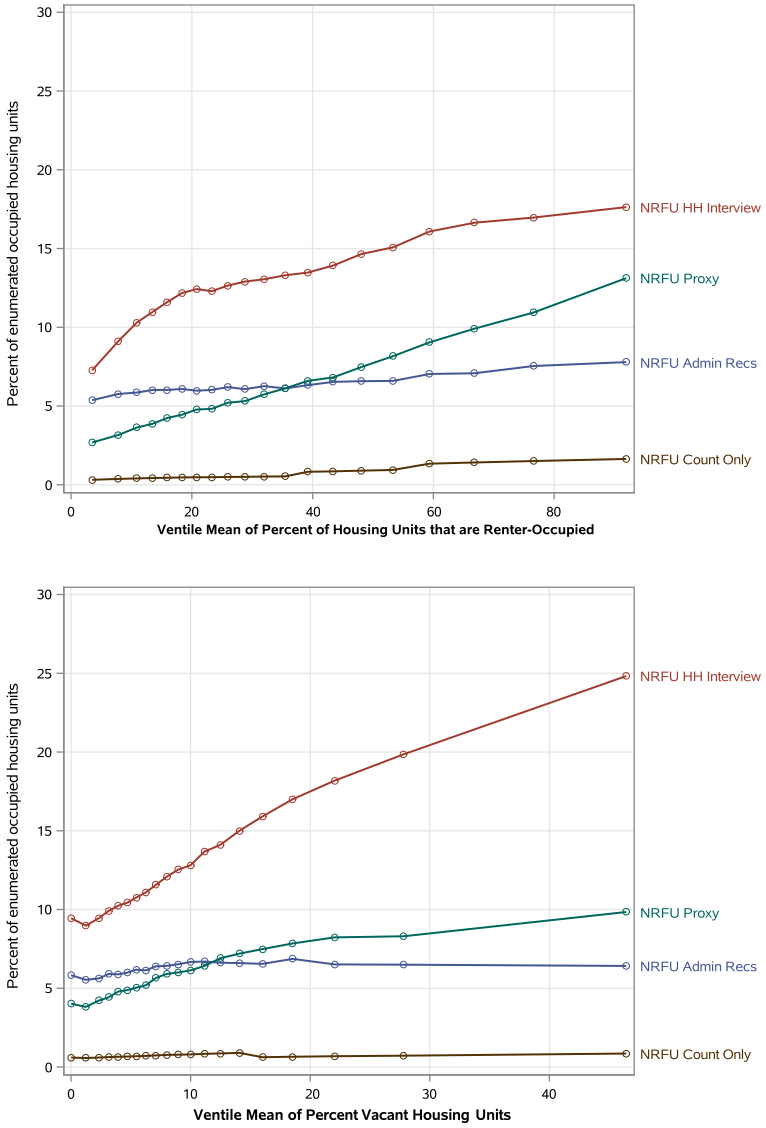


Figure 7.16 Mean percent Nonresponse Followup resolutions by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent housing units renter occupied and (b) percent vacant housing units.

NOTES: NRFU, Nonresponse Followup; HH, householder.

SOURCE: NRFU Resolutions by Phase Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0224.

Copyright National Academy of Sciences. All rights reserved.

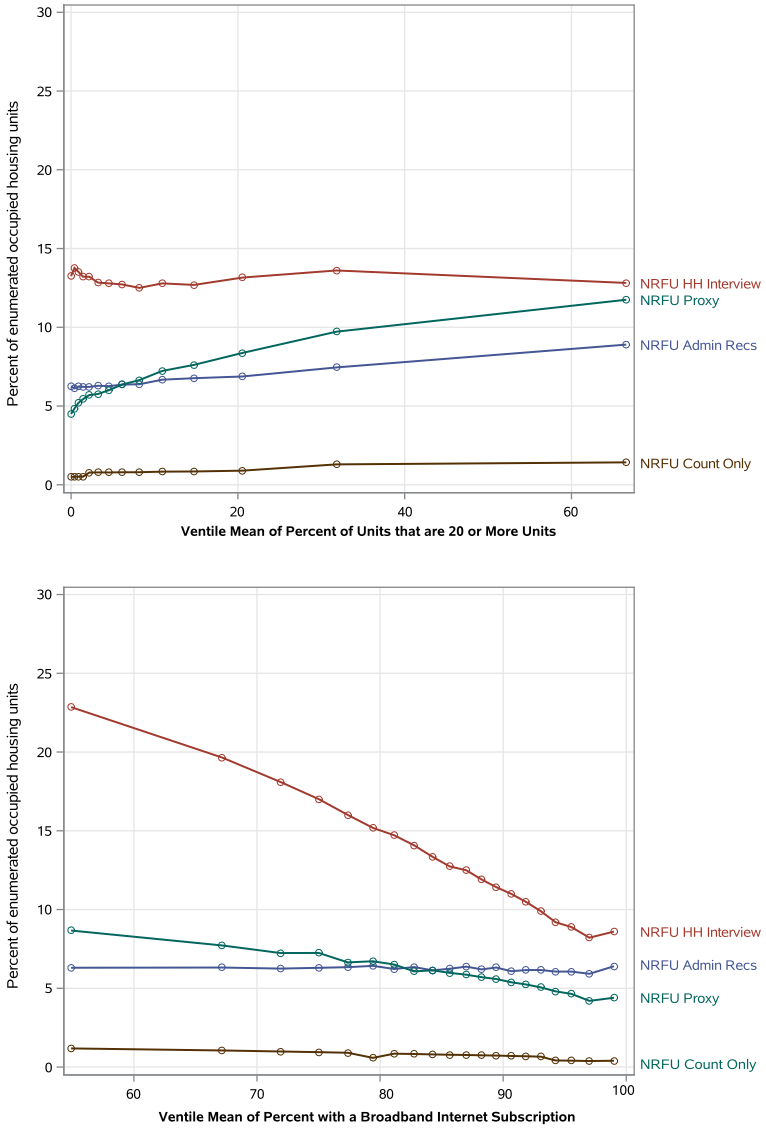


Figure 7.17 Mean percent Nonresponse Followup resolutions by mode, grouping census tracts by ventiles of American Community Survey characteristics: (a) percent structures with 20 or more units and (b) percent with broadband internet subscription.

NOTES: NRFU, Nonresponse Followup; HH, householder.

SOURCE: NRFU Resolutions by Phase Dataset and select 2016–2020 American Community Survey characteristics. See Disclosure Review Statement; CBDRB-FY23-0224.

Copyright National Academy of Sciences. All rights reserved.

attainment and with lower household income appear to be represented in the 2020 Census with poorer-quality data. This pattern was similar for census tracts with higher unemployment rates. Tracts with high unemployment were generally more likely to be enumerated by NRFU, with a higher proportion enumerated by household interview, proxy, and count only compared with census tracts with lower unemployment rates.

Census tracts with high percentages of non-Hispanic White people were less likely to be enumerated in NRFU (Figures 7.14–7.15). Those census tracts that were over 90% non-Hispanic White were more likely to be enumerated by household interviews compared with those census tracts that were between 60–90% non-Hispanic White. This is likely because census tracts that are 90% or more non-Hispanic White are generally rural census tracts that tended to be Update Leave areas. The patterns of NRFU modalities for percent non-Hispanic Black people, Hispanic people, and people speaking a language other than English were very similar and tended to be the opposite of patterns for non-Hispanic White people. Generally, the quality of 2020 Census data representing areas with concentrations of African American people and Hispanic people, and areas with populations that do not speak English well tend to have lower incomes.

Higher concentrations of renter-occupied housing were associated with higher enumeration by NRFU (Figures 7.16–7.17). An average of more than 12% of housing units were enumerated by proxy in areas where approximately 90% of housing was renter occupied. Census tracts with more vacant housing units were also more likely to be enumerated in NRFU. Census tracts with about 45% of vacant housing units had higher percentages of households enumerated by household interviews and proxy interviews. Census tracts with higher percentages of housing units being housing complexes with 20 or more units appeared more likely to be enumerated by proxy interview compared with census tracts with smaller proportions of housing being large multifamily complexes. The percentage of households with broadband internet subscriptions was negatively associated with enumeration by NRFU. Census tracts with smaller proportions having broadband were more likely to be enumerated by household and proxy interviews.

7.5 CONCLUSIONS

Just as it is difficult to imagine successful completion of 2020 Census data collection without Internet Self-Response, so too is it difficult to imagine any successful 2020 Census data collection if the traditionally resource-intensive NRFU relied on physical transfer of assigned questionnaires/cases and daily face-to-face meetings among data-collection crews, as was the case in 2010. In late stages of 2020 Census planning, the Census Bureau partially pulled back

Copyright National Academy of Sciences. All rights reserved.

from full automation of workload assignments among the enumerator employee pool, instead planning for a "less is more" approach in three phases of NRFU, putting the last (and arguably most difficult) 15% of NRFU cases in specific areas in the hands of a reduced set of enumerators—those who had proven most productive in the first phase. The COVID-19 pandemic-related replanning and the shifting target date for completion of data collection substantially impacted NRFU; the Census Bureau was legally barred by court ruling from dismissing field staff (and thus from conducting the new, phased approach to NRFU exactly as planned). The 2020 Census NRFU operation was also very different from its predecessors due to its use of administrative records data, not only to flag potential vacant and delete-worthy cases early on (for removal from fieldwork after one confirmatory field visit) but as a means for enumerating households if information of sufficient quality existed for a nonresponding household.

NRFU is an inherently challenging census operation from the quality standpoint, and the Census Bureau routinely pairs NRFU with an aggressive quality-assurance operation including reinterview by different enumerators. Recall bias, or respondent difficulty in providing accurate information about a fixed (past) reference date, is always a concern, as are opportunities for miscoding and data-entry error by enumerator staff and the perennial worries about "curbstoning" (deliberate fabrication of reported information to complete cases quickly). All of these factors, and more, had particular salience in the 2020 Census with the COVID-19 pandemic-forced delays and constrained timeline; the performance of data-entry systems and safeguards on enumerator smartphones, new in application in the 2020 Census, also undoubtedly had some effect on NRFU operations. Amidst this complex set of issues, our analysis was particularly constrained by the lack of evaluation results from the Census Bureau itself, if only to provide us with a better sense of the operational paradata actually retained during the NRFU operation; so we were compelled to keep our queries at a fairly high level. The original intent of the panel's data analysis subgroup was to try to assess the general question of whether 2020 Census data collection was unduly rushed, by examining data on the number and timing of enumerator visits to NRFU case households. However, in discussion with Census Bureau staff, we concluded that it would be too difficult and disruptive to extract those data from the broader Census Data Lake that underlies the 2020 Census information technology architecture—given that the Census Bureau's own researchers had not yet compiled such an evaluation dataset for their own reports. Accordingly, as a compromise measure, we focused on available data on the resolution of NRFU cases, broken down by the three defined phases of NRFU for 2020.

Conclusion 7.1: Analysis of the resolution of Nonresponse Followup (NRFU) cases by operational phase does not, in itself, show signs

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

that 2020 Census data collection was unduly hastened to completion. Enumeration by proxy report or by administrative records (if records information of sufficient quality was available) was invoked in all stages of NRFU, but not in a way suggestive of an egregious rush to closeout.

Conclusion 7.2: Though the logic of the U.S. Census Bureau's split of Nonresponse Followup (NRFU) into distinct operational phases is generally sound, it is impossible to assess the effectiveness of the phased approach as originally designed because the circumstances of the 2020 Census precluded the planned winnowing of NRFU staff. The determination of change from one NRFU phase to another was data driven, based on resolution of a specified percentage of cases in a particular geographic area, which is appropriate. However, the specific operational geography at which NRFU phases were determined—Census Field Supervisor Areas—was so operational and tactical in nature, and so divorced from familiar tabulation geography, that it is unworkable from an evaluation standpoint and raises concerns that it might be difficult for census managers to appreciate actual completion coverage during live field operations.

Conclusion 7.3: The information available for the panel's analysis did not permit assessment of the potential effects of the delays in 2020 Nonresponse Followup (NRFU) and other field operations in 2020 on the quality of results, such as recall biases and the coding of accurate occupancy statuses of dwellings (especially later in the NRFU period).

The analyses presented illustrate significant problems with variation in the quality of 2020 Census data for areas characterized by some indicators of lower socioeconomic status. Further assessment of the impact of response modes on data quality as associated with demographic, economic, and housing composition is suggested. The results presented do not attempt to assess the magnitude of the data-quality issues. More research is needed on the quality of data generated in the imputation process. Results from such research could be employed to estimate the magnitude and impact of imputation for item nonresponse.

Not all proxy enumerations are the same. In some cases, proxy informants may be very familiar with residents of the reference housing unit and provide accurate responses regarding the number of residents, residents' race and ethnicity, sex, and perhaps even age. Yet, at the other extreme, proxy respondents may know little about the residents of the reference housing unit. In densely populated areas with high resident turnover, complex family arrangements, or a range of other challenges and combinations of challenges to knowing neighbors, the quality of proxy responses will more likely be less

Copyright National Academy of Sciences. All rights reserved.

than accurate. We assume that proxy and count-only enumerations result in substantial item nonresponse and consequently in poorer-quality data.

Conclusion 7.4: Nonresponse Followup (NRFU) enumerations result in more item nonresponse and poorer coverage than self-responses. NRFU item nonresponse for race and Hispanic origin exceeded 15% in the 2020 Census. Proxy enumerations had greater than 35% item nonresponse for race and Hispanic origin. Results suggest that census tracts with higher proportions of people with lower socioeconomic status (SES) characteristics were more likely to be enumerated using modes that resulted in higher item nonresponse. The net result is that 2020 Census data on race, Hispanic origin, and age are less accurate in those census tracts than in those with higher proportions of higher SES characteristics. This result is most relevant for redistricting. State and local officials are defining political districts with data that vary in how accurately they reflect the age and racial and ethnic composition of the population. Census data are particularly prone to error in areas with high percentages of people who are Black, Hispanic, or American Indian and Alaska Native living on reservations.

Recommendation 7.1: As a goal for the 2030 Census, the U.S. Census Bureau should consider a major reduction in the use of proxy interviewing for enumeration, if not the elimination of proxy reporting in all but very limited circumstances. Work toward this goal should be predicated on the results of research that only the Census Bureau can conduct, reexamining and comparing proxy reports in the 2020 Census with the information that would have been available from administrative records or third-party data resources.

Recommendation 7.2: The U.S. Census Bureau should consider conducting research on the quality, and factors associated with variation in the quality, of proxy and count-only responses, with the aim of minimizing the occurrence of these factors.

To repeat an observation from Chapter 4, errors in the census, whether coverage or characteristics, would matter less if they were equally distributed among population groups and across geographies. Yet there is ample evidence (see Chapters 6 and 7) that self-response is not uniform across the population and therefore that NRFU is not uniform. For example, census tracts with higher percentages of renters and Hispanic households and lower median incomes based on the ACS had higher rates of NRFU enumerations, and of proxy NRFU enumerations in particular, in 2020 compared with other tracts.

These patterns strongly suggest that there are inequities among population groups and geographies in the completeness and quality of 2020 Census data.

Copyright National Academy of Sciences. All rights reserved.

Conducting NRFU for millions of addresses under the pressure of the census time schedule is challenging, and it is to the Census Bureau's credit that it does so under difficult circumstances. It is possible, moreover, that the quality of the 2030 NRFU will revert to the historical mean and that 2020 was an aberration due to the COVID-19 pandemic. It is also probable, given declining response rates to surveys around the world (including item as well as household nonresponse), that it will become increasingly difficult over time to enumerate the population in NRFU. Administrative records may help, although they are not a panacea—for example, they are often missing race and ethnicity and are not available equally across geographic areas and population groups (see Chapter 8).

Other chapters discuss needed developments and testing for methods to improve self-response and the use of administrative records for enumeration, which would reduce the extent of the NRFU operation. For the 2030 Census, however, it is likely that a significant NRFU component will be needed. A prerequisite for improving the quality of the data obtained in NRFU is for the Census Bureau to conduct research at a fine-grained scale on the geographic and population distribution of NRFU and particularly of proxy NRFU responses; we formally recommend the same, in line with our other recommendations in Chapter 12. Coupled with the research program outlined in Appendix D, the results from fine-grained analysis of NRFU could position the Census Bureau to develop specialized approaches for improving NRFU in 2030 for particular areas and groups. Publishing detailed results of its NRFU analysis would also help census data users assess the relative quality of the 2020 Census for specific areas and groups, and would contribute evidence-based suggestions for improving the 2030 Census.

Recommendation 7.3: The U.S. Census Bureau should conduct fine-grained analysis of Nonresponse Followup enumerations (separately for household, proxy, and administrative records responses) for small geographic areas and population groups as part of its research to develop and test ways to improve the quality of enumeration in the 2030 Census. The Census Bureau should publish the results of its analysis as an aid to users of 2020 Census data.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

– 8 –

Use of Administrative Records for Enumeration in the 2020 Census

Administrative records data are typically collected as a necessary step in the administration of governmental programs. Key U.S. examples include personal income tax data and Medicare data. The U.S. Census Bureau has used administrative records for decades in its census and survey programs. It produces annual population estimates by updating a prior census with birth, death, and other records (see Chapter 10). In turn, it uses the population estimates to reweight surveys to address differential undercoverage by age, sex, race, and ethnicity. Demographic Analysis, one of the major methods of census coverage evaluation, uses the same kinds of records as the population estimates to produce independent estimates of the population (see Chapter 4). The Census Bureau uses tax records for small businesses to eliminate the need to survey these businesses for the economic census,[1] and it has long used records from the U.S. Postal Service (USPS) to update its Master Address File (MAF) (see Chapter 5).

With rising costs of field data collection, statistical agencies in the United States and abroad have sought new and expanded ways to use administrative records for censuses and surveys. These ideas include using administrative records: (1) to help in the development of a survey or census frame; (2) for editing and imputation of missing survey and census responses; (3) for survey or census evaluation; (4) to improve the operational efficiency of survey or census

[1]See the technical documentation for the economic census at https://www.census.gov/programs-surveys/economic-census/year/2022/technical-documentation.html.

Copyright National Academy of Sciences. All rights reserved.

data collection; or (5) as a replacement for survey or census data collection for all or a subset of the units (see relevant discussions in National Academies of Sciences, Engineering, and Medicine, 2017, and National Research Council, 1994).

The idea of using information from administrative records to enumerate the U.S. population may have first been raised over 40 years ago by Alvey and Scheuren (1982). To date, the notion of conducting a census of the United States by using only the data collected in administrative records, as is done in several countries, has not garnered support. Reasons include records' unequal coverage of all population groups and their frequent lack of such characteristics as race and ethnicity. Nonetheless, because of the cost advantages and, for some purposes, quality advantages (e.g., for estimating various types of income in surveys), the Census Bureau has invested considerable effort over the past two decades to expand, evaluate, and conduct research on applications of available administrative records. Research leading up to the 2020 Census focused on using administrative records information to replace some of the field enumeration conducted during Nonresponse Followup (NRFU). Specifically, the idea was to use information available from administrative records as a form of census enumeration in a limited supplementary role, in cases for which an initial contact did not elicit a response and the Census Bureau felt it had sufficiently high-quality administrative records for the household that could be used. The Census Bureau also planned research to compare a compilation of administrative records with actual 2020 Census results.

In this chapter, we focus on the first-ever use of administrative records data in a U.S. census for enumeration of selected households that did not self-respond and for which one enumerator visit in NRFU also failed to produce a return. The hope was that NRFU data-collection costs would be reduced, while the information quality would be equal or better using administrative records instead of further NRFU field work. As it turned out, 3.8% of occupied units were enumerated with administrative records, either as planned after the first attempt at a NRFU interview or in NRFU closeout (see Chapter 7 for discussion of NRFU). In addition, 14.4% of vacant units were enumerated with administrative records, as were 3.4% of addresses that turned out to be nonresidential or nonexistent. In total, 4.6% of all addresses were resolved using administrative records (calculated from U.S. Census Bureau, 2021a). At the end of the chapter, we briefly consider the possible role of administrative records enumerations in the 2030 Census. For the role that administrative records played in enumeration of various group quarters populations in 2020, see Chapter 9.

Copyright National Academy of Sciences. All rights reserved.

8.1 ADMINISTRATIVE RECORDS ENUMERATIONS IN THE 2020 CENSUS

Using administrative records for enumeration after a failed NRFU visit may seem straightforward in concept, but there are a number of complexities that make implementation a challenge. First, the information contained in administrative records data can pertain to occupancy of a residence by different groups of people at different time periods close to Census Day. Second, the residence rules of the decennial census may differ from residence rules employed by various governmental programs whose administrative records might be considered for enumeration. Third, the Census Bureau's use of a composite of national administrative records databases requires reconciling differences in information for the same address/household across individual records sources. Fourth, the definition of "household members" can vary depending on the governmental program. For instance, the residents of a household on personal income tax forms, according to the Internal Revenue Service (IRS), can be a subset of the residents of a household as defined by the Census Bureau, since the IRS does not include nondependents who are not members of the nuclear family. Also, not all children qualify as dependents, and children and other relatives do not have to be members of the taxpayer's household all year to be claimed as dependents.[2] Finally, the quality of some administrative records can be poor, especially with respect to the characteristics information that the decennial census collects. Poor quality involves not only missing information on race and ethnicity in many instances, but also some amount of missing or erroneous entries for other characteristics.

Before discussing the statistical models used for administrative records enumeration in select circumstances, it is important to point out the effects of the COVID-19 pandemic on the operation (see Appendix C.7.5). COVID-19 delayed the collection and delivery of some of the administrative records files that various federal agencies had agreed to provide. In particular, it delayed receipt of IRS data because of postponed deadlines for the filing of personal income tax forms. To err on the conservative side, the Census Bureau also scaled back its plans to use administrative records to determine that addresses were vacant or nonresidential. On the other hand, the pandemic motivated the Census Bureau to make an unplanned use of administrative

[2]From Mulry et al. (2023b): "The IRS has rules about whom a taxpayer can claim as a dependent on a 1040 tax return. There are two basic sets of rules for when a taxpayer may claim a person as a dependent, one set applies to children and the other applies to other relatives. There are situations when a child does not qualify as a dependent child but does qualify as a dependent relative. Usually, a dependent must reside with the household of the taxpayer all year, but there are situations where a close relative does not have to reside with the household. However, the taxpayer does have to provide support for the dependent for at least half of the year. ... To meet the Qualifying Child Test, the child must be ... either younger than 19 years old or be a student younger than 24 years old as of the end of the calendar year."

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

records for the closeout phase of enumeration, allowing enumeration based on one administrative records source rather than multiple sources.

8.2 MODELS DEVELOPED FOR THIS PURPOSE

Deciding to use administrative records information to enumerate an address in the NRFU workload involved three decision points, each supported by probabilistic models. The first decision was to estimate the probabilities that the address was not a housing unit (either because it was nonresidential or nonexistent), that it was a vacant housing unit, and that it was an occupied housing unit. The second decision was to estimate the probability that, if occupied, the number of residents was equal to various possible counts. The third decision was to judge whether the quality of the information used for the first two decisions was sufficiently high to rely on it for census enumeration, or instead to return the enumeration of the housing unit to NRFU.

The Census Bureau used information from multiple administrative records databases as input to these models, each set of records having advantages over the others, especially for specific subpopulations. The primary databases used in 2020 for the decision models were: IRS 1040 forms filed for 2019, IRS 1099 forms filed for 2019, Medicare records, the Indian Health Service Patient Database, and the Household Composition Key File at the U.S. Census Bureau.[3] In addition, a number of databases could be used solely for corroboration of the information found in the above databases: USPS Change of Address File, 2000 and 2010 Census Unedited File, Medicare files, U.S. Department of Housing and Urban Development (HUD) Public and Indian Housing Information Center, HUD Tenant Rental Assistance files, HUD Computerized Home Underwriting Management System, HUD Public and Indian Rental Certification System, Selective Service Registration file, Targus Wireless, Targus Federal Consumer, Targus Address file, Veteran Service Group of Illinois Name and Address Resource file, Veteran Service Group of Illinois Tracker Plus, and Data, Analytics, and Research Partners.

Linking the information from these disparate sources was a key step. To enable such linkage, the Census Bureau assigned an anonymized identifier to each person record in each administrative records system, using the Census Bureau's Person Identification Validation System (PVS). The PVS determines a protected identification key (PIK) via a probabilistic matching algorithm

[3]The Household Composition Key File is a database created by Census Bureau staff using applications for Social Security numbers (SSNs) from the Social Security Numerical Identification (Numident) File, which contains names and SSNs that are used in assigning protected identification keys (PIKs) to records, enabling linkage of child-to-parent records. See the Spring 2017 Administrative Records Modeling Update for the Census Scientific Advisory Committee for further information, at https://www2.census.gov/cac/sac/meetings/2017-03/admin-records-modeling.pdf.

Copyright National Academy of Sciences. All rights reserved.

between the administrative record source data and a series of reference files. The algorithm used for determining administrative records enumerations assumed that the PIK assignment was correct, though it is important to acknowledge the possibility of linkage error associated with the PVS methodology. (See Layne et al., 2014, and Wagner and Layne, 2014, for details on PVS methodology.) The Census Bureau then used these linkable administrative records files to create household rosters for all housing units in the decennial census universe. These rosters are the collection of PIKs found in any of the administrative records files at a given address. Characteristics of the housing unit and of its occupants are provided by sources such as the Social Security Numerical Identification (Numident) File and the most recent decennial census. It is important to note that the algorithm could not proceed unless the housing unit existed on the MAF and therefore had a MAFID. Children were linked to rosters for each of the addresses where either the mother or father appeared.

In thinking of these models, it is important to stress that they are premised on the existence of undercoverage and overcoverage between census returns and administrative records data, and on understanding the properties of those groups. That is, there should be substantial overlap between households/persons in the census and households/persons in the composite of administrative records sources formed through linkage. But there should be important differences as well, with some households/persons appearing in administrative records but not in the census, and vice versa. Finally, it is important to note that the models draw predictor variables from a variety of sources as well: the USPS Delivery Sequence File, the American Community Survey (ACS), the MAF, census operational information, and USPS Undeliverable as Addressed (UAA) reason codes.

8.2.1 Housing Unit Status Model

To discriminate between nonexistent or nonresidential (deleted), vacant, and occupied housing units, the Census Bureau developed a multinominal logit model using housing unit-level administrative records data as regressors to predict these three possibilities for housing unit status. The indicator variables denoting housing unit status are:

$$y_h^{\text{unocc}} = \begin{cases} 1, & \text{if occupied} \\ 2, & \text{if vacant} \\ 3, & \text{if deleted} \end{cases}$$

The "unocc" superscript denotes the model used for administrative records-based determination of housing status. The subscript h indexes the individual housing units. The multinominal logit model is used to estimate the probability of each status type for each housing unit. The predicted probabilities denoted

Copyright National Academy of Sciences. All rights reserved.

by:

$$\begin{aligned}
\hat{p}^{\text{unocc}}_{h,\text{occ}} &= P\left(y^{\text{unocc}}_{h} = 1\right), \\
\hat{p}^{\text{unocc}}_{h,\text{vac}} &= P\left(y^{\text{unocc}}_{h} = 2\right), \\
\hat{p}^{\text{unocc}}_{h,\text{del}} &= P\left(y^{\text{unocc}}_{h} = 3\right)
\end{aligned}$$

are passed to the distance function, described below, to determine which cases are identified as being of each type (or are left undecided). The covariates for this model include the following ACS variables (all as percentages at the block-group level):

- Persons in block group between 25 and 44;
- Persons in block group greater than 64;
- Persons in block group identifying as Black;
- Persons in block group identifying as Hispanic;
- Occupied housing units in block group with at least two related household members;
- Persons over age 4 in block group speaking language other than English at home;
- Housing units in block group considered as mobile homes;
- Housing units in block group where householder/spouse are members of household;
- Occupied housing units in block group that are not owner occupied;
- Housing units in block group occupied at time of interview; and
- Persons in block group living below the poverty level.

The covariates for this model also include the following individual housing unit characteristics:

- Number of neighbors in NRFU;
- USPS UAA reason; and
- Housing unit type (e.g., multifamily).

Housing unit variables collected from administrative records for inclusion in the model include flags indicating whether more than one person in the household falls into any of the following categories: Black; Hispanic; missing ethnicity; and less than 2, 2–10, 10–17, 18–24, 25–44, and over 64 in age. Finally, there are indicator variables for which of the various administrative records information sources are available.

8.2.2 Models to Estimate the Number of Occupants

Once the probability of housing unit status was modeled, the Census Bureau also needed to model the probability that various numbers of occupants

Copyright National Academy of Sciences. All rights reserved.

were resident in the occupied housing units. The Census Bureau defined seven valid household compositions for estimation purposes, six defined as the combination of {1 adult, 2 adults, or 3 adults} with {0 children or at least 1 child} and the seventh being {at least 4 adults}. With this base, the Census Bureau developed two competing models to estimate the number of occupants, referring to them as the person-place model and the household composition model.

8.2.3 Person-Place Model

The person-place model estimated the probability that a person found in the union of the administrative records files at an address would have been enumerated at that same address in the 2020 Census had the NRFU operation been completed using field work. The Census Bureau fit a logistic regression at the person level using 2010 Census data to predict the following outcome:

$$y_{ih}^{\text{occ1}} = \begin{cases} 1, & \text{if person } i \text{ is found in administrative records} \\ & \quad \text{and 2020 Census at same address} \\ 0, & \text{otherwise} \end{cases}$$

where the occ1 superscript denotes that the housing unit status model predicted that the housing unit was occupied, and that, given this, the person-place model was used to provide the probability that a given individual was resident, with h denoting the address in question and i denoting the administrative records person. For all person-place pairs in the union of the administrative records files, this model provides a predicted probability that the 2020 Census and the administrative records roster data place the person at the same address, and is denoted:

$$\hat{p}_{ih}^{\text{occ1}} = P\left(y_{ih}^{\text{occ1}} = 1\right)$$

The 2020 Census person records are assigned PIKs with the methodology summarized above, which are used to identify duplicates across administrative records sources. The predictor variables used to estimate the above probability include: indicators of the presence of the administrative records person in each source at the address; indicators of the presence of the administrative records person at a different address within the same administrative records source; address-level administrative records information (e.g., the number of administrative records people associated with an address); field operations information (e.g., USPS mailing information, number of NRFU neighbors); and information from other survey sources (e.g., characteristics of the local geography, such as poverty rate, renter rate, and vacancy rate).

While the person-place model is fit at the person level, the decision on the housing unit is made at the housing unit level. Therefore, the person-level

Copyright National Academy of Sciences. All rights reserved.

predicted probabilities must be combined to provide the probability that the household is occupied with the same number of people as the census would have counted, as follows:

$$\hat{p}_h^{\text{occ1}} = \min\left(\hat{p}_{1h}^{\text{occ1}}, \ldots, \hat{p}_{n_h h}^{\text{occ1}}\right).$$

This use of the minimum estimated probability for any number of occupants between 1 and n_h (the maximum number of persons the administrative records have "shown" to be resident at the address in question) is acknowledged by the Census Bureau as a conservative estimated probability. The Census Bureau justified this conservative approach given that this type of enumeration with administrative records had not been attempted previously.

8.2.4 Household Composition Model

The household composition model is used to estimate the probability that the address in question has the same household composition (number of adults and children) determined through searches through the administrative records databases, as would have been determined by NRFU fieldwork (had it been conducted for that address). The Census Bureau fit a multinomial logit model on 2010 housing unit-level administrative records data and 2010 Census data and used those estimated coefficients to predict the outcome from the 2020 Census. The dependent variable had the following form:

$$y_h^{\text{occ2}} = \begin{cases} 0, & \text{if unit } h \text{ is vacant in the 2020 Census} \\ 1, & \text{if unit } h \text{ has 1 adult and 0 children in the 2020 Census} \\ 2, & \text{if unit } h \text{ has 1 adult and at least 1 child in the 2020 Census} \\ 3, & \text{if unit } h \text{ has 2 adults and 0 children in the 2020 Census} \\ 4, & \text{if unit } h \text{ has 2 adults and at least 1 child in the 2020 Census} \\ 5, & \text{if unit } h \text{ has 3 adults and 0 children in the 2020 Census} \\ 6, & \text{if unit } h \text{ has 3 adults and at least 1 child in the 2020 Census} \\ 7, & \text{if unit } h \text{ has at least 4 adults in the 2020 Census} \end{cases}$$

The occ2 superscript denotes that the household was predicted to be occupied by the housing unit status model and that the household composition model was being used. The h subscript denotes the individual housing unit. For every address in the NRFU universe, this model provides the predicted probability of each of the eight household composition types: $\hat{p}_{h,k}^{\text{occ2}} = P\left(y_h^{\text{occ2}} = k\right)$, for $k = 0, 1, \ldots 7$ (see above). The covariates for this multinomial logit model include those from housing unit-level administrative records information (i.e., the count of all administrative person records associated with the address), person-level administrative records information (e.g., indicators of whether any

Copyright National Academy of Sciences. All rights reserved.

administrative records person attached to the address in question was found at a different address within the same administrative records source), and housing unit-level information from other survey sources (e.g., flags indicating that young children were associated with the household). The Census Bureau then conditioned on the household composition that was determined by searches through the relevant administrative records databases to provide a k^*, the best administrative records composition. Then, $\hat{p}_h^{\text{occ2}} = \hat{p}_{h,k^*}^{\text{occ2}}$ is the probability of the household composition sent forward from the household composition model. For example, $\hat{p}_h^{\text{occ2}} = \hat{p}_{h,3}^{\text{occ2}}$ would be the estimated probability for a housing unit with an administrative records household composition type of two adults and no children. These predicted probabilities are input into the distance function (discussed below) to determine which cases are identified as "AR Occupied."

8.3 JOINING THE THREE MODELS—DISTANCE FUNCTIONS

The initial idea considered by the Census Bureau was to combine information from the housing unit status, person-place, and household composition models using linear programming (LP) techniques to determine the occupancy status of the housing unit. The possible status codes for the housing units were nonresident or nonexistent, vacant, or occupied. For those housing units that were occupied, the number of people that were residents was represented in the LP models. For example, to identify administrative records-estimated vacant units, this method set linear constraints that the average vacant predicted probability must exceed a prespecified threshold. In addition, the sum of the predicted probabilities for occupied units could not exceed a certain percentage of the estimate of occupied housing units from the ACS. For identifying administrative records-based estimated number of residents for occupied units, the constraints were that the average person-place predicted probability must exceed a prespecified threshold, and the average household composition-predicted probability must exceed a different prespecified threshold. Without getting into specifics, this approach ran into problems setting the various thresholds due to: (1) the fact that the averaging performed allowed the estimated probabilities for other households to contribute to the identification of a household as being vacant; and (2) the geography over which the averaging was done mattered a great deal.

The above approach was replaced by an alternative approach using a distance function, which overcame or reduced these concerns. Each housing unit was evaluated on its own, and the thresholds used were more transparent and interpretable. This distance function method was also easier to implement in real time. Starting with the predicted probabilities for vacant and occupied, $\hat{p}_{h,\text{vac}}^{\text{unocc}}$ and $\hat{p}_{h,\text{occ}}^{\text{unocc}}$, as estimated by the housing unit status model, these two probabilities can be thought of as points on a two-dimensional plane of

Copyright National Academy of Sciences. All rights reserved.

estimated probabilities. The most likely vacant cases would be those that have the shortest distance to the point at which the estimated occupied probability equals 0 and the estimated vacant probability equals 1. The *vacant* Euclidean distance from this "perfect" point of estimated probabilities is then d_h^{vac}, where for each unit h, $d_h^{\text{vac}} = \sqrt{\left(1 - \hat{p}_{h,\text{vac}}^{\text{unocc}}\right)^2 + \left(\hat{p}_{h,\text{occ}}^{\text{unocc}}\right)^2}$.

With regard to enumerating occupied housing units using administrative records, the Census Bureau started with the predicted probabilities of number occupied, $\hat{p}_h^{\text{occ1}}$ and $\hat{p}_h^{\text{occ2}}$, from the two models. These can be thought of as two measures of quality (probability of count match and probability of household composition match, respectively), such that the housing units with higher-quality administrative records are associated with higher estimated probabilities. Even though the predictions from these two models are clearly correlated, research showed that important additional information was gained by using the estimated probabilities from both models. Here, the ideal point in the two-dimensional plane of estimated probabilities is $(1, 1)$, indicating an estimated probability of 1 that the administrative records count will equal the decennial census count, and likewise an estimated probability of 1 that the administrative records household composition will equal that of the decennial census composition. The corresponding distance function here is:

$$d_h^{occ} = \sqrt{\left(1 - \hat{p}_h^{\text{occ1}}\right)^2 + \left(1 - \hat{p}_h^{\text{occ2}}\right)^2}.$$

When implemented, two cut-off thresholds are selected for d_h^{vac} and d_h^{occ} to determine housing unit status and number of residents. When those thresholds are satisfied, they become the enumerations. When those thresholds are not satisfied, the unit is returned to NRFU because the administrative records information has been judged as being of insufficient quality.

There are three remaining issues regarding implementation. The first issue is the mechanism for providing characteristics information for the people enumerated in this manner. Administrative records information is generally available for sex and age. Information for race and ethnicity is somewhat less available, though for many cases, one can collect information from the Numident file or from the previous census.

The second issue is how to estimate the coefficients of these three models. One needs information on how households were enumerated by the decennial census data collection in the field and the relevant information for each household from the existing administrative records databases. In other words, a training set is needed. In 2020, data were used from the 2010 Census, with the assumption that the estimated coefficients remained relevant over the decade. If that assumption seemed too strong, one could use data from a large census test that included a NRFU component.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Third, one specific problem that will arise in using this approach in 2030 is that no field data collection was conducted for those housing units that were enumerated using administrative records and therefore were not returned to NRFU in 2020. In other words, there will not be a training set relevant to the entire population. For that reason, it might be beneficial to follow through with NRFU for at least a sample, or maybe for all cases in a large census test to find out whether things have changed since 2020.

8.4 EVALUATING THE APPROACH USED IN 2020

In the 2020 Census, the Census Bureau enumerated about 4.6% of the verified MAF addresses using administrative records in this first-of-its-kind application. It is useful to point out that a number of groups, including previous National Academies of Sciences, Engineering, and Medicine Committee on National Statistics panels, had recommended that the Census Bureau research the use of administrative records for this purpose for at least two censuses, and the fact that this method was used to enumerate a nontrivial fraction of the country should be regarded as a success.

However, this operation was conducted in conjunction with the promise that, in doing so, a considerable amount of public funds would be saved (possibly to be used in other aspects of the taking of the decennial census) and that the quality of the resulting counts would be at least as good as if NRFU had been allowed to proceed for its complete workload. Given that 4.6% of all the enumerations represents about one-seventh of the NRFU universe, this suggests that hundreds of millions of dollars may have been saved. However, it should be cautioned that the savings from this use of administrative records enumeration may have been less than anticipated because the temporary field staff levels were affected by shifting schedules and court orders. To our knowledge, no data on actual or potential cost savings are available.

Given the Census Bureau's general perception that the quality of the resulting enumerations was relatively high, an expanded version of this application will likely be used again in the 2030 Census. Therefore, it is important to examine what is currently known about the quality of these enumerations and how more can be learned prior to 2030. At a minimum, the Census Bureau needs to evaluate the general quality of the resulting counts and the success of the three models used for this purpose (as well as the distance functions)—both to understand their strengths and weaknesses and to examine the possibility of replacing them with superior alternatives, if any.

The Census Bureau has conducted some evaluations of this recent use of administrative records, both in census tests before the 2020 Census and in the 2020 Census itself. Focusing on the procedure used in 2020, Mulry et al. (2023a,b) are the primary reports available that assess the 2020 administrative

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

records enumeration. The main analysis tool contributing to these evaluation(s) of the use of administrative records for enumeration of some NRFU households in 2020 was a comparison of rosters for self-responding households with the corresponding rosters for the housing unit from administrative records, and similarly, rosters for NRFU respondents (excluding proxy respondents and units requiring whole-household imputations) compared again to the corresponding administrative records rosters. This was done with the aim of developing a better understanding of the degree of agreement between rosters developed from administrative records and rosters collected in the census.

Given this approach, these evaluations were not intended as a direct assessment of the use of administrative records in 2020 for enumeration of some NRFU households. This is evident in a variety of ways. First, the population for which the Census Bureau attempted an administrative records enumeration during the 2020 Census consisted of those households that did not respond to the first attempt at a NRFU enumeration. Therefore, while both the evaluation study and what transpired during the 2020 Census examined those households personally interviewed as part of NRFU for the second through sixth interview attempts, the evaluation also included both self-respondents and those who were successfully interviewed for the first NRFU interview. The 2020 Census implementation of this type of enumeration also included cases that would have been proxy cases and whole household imputations had they not been removed from NRFU. Second, these evaluations do not address other implementation issues, such as the use of administrative records for enumerating NRFU closeout cases through use of specific thresholds. In that application, the thresholds were lowered to relax the various quality standards in place for noncloseout NRFU cases, and corroborating administrative records were not required. The evaluation reports calculated that, of those enumerated using administrative records, 92% of households were of the typical situation—only one failed NRFU visit, with corroborative administrative records available—rather than close-out cases. In addition to checking competing rosters for their degree of agreement, these evaluations examined the impact of both the mode of response (Self-Response, NRFU) and the timing of the enumeration on the degree of agreement, as well as the impact of the demographic makeup of the household.

Summarizing the findings in Mulry et al. (2023a,b), nationally, of the 152 million occupied housing units, 66 million (43%) had a high-quality administrative records-based roster for use in enumeration. Of the NRFU population's 16.5 million housing units (which includes those who were interviewed on a first visit and therefore were excluded from consideration for this type of enumeration), only 3.8 million (23%) had high-quality administrative records-based rosters for use in enumeration. Therefore, one might be concerned that, unless some substantial changes are made in the

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

quality of administrative records by 2029, this method could be difficult to implement for a majority of the NRFU population in 2030.

For Self-Response households having both a census roster and a high-quality administrative records roster, about 79.5% had population counts that were in agreement; the difference between census and administrative records counts for these Self-Response households was +1 for 8.6% and −1 for 8.1%. Hence, the percentage of Self-Response households with rosters in which the counts were at worst one person apart was 96%. However, for NRFU cases—more relevant to the application of enumeration in the 2020 Census—the percentage of exact agreement dropped to 58.7%. An additional 15.8% and 12.7% of NRFU households differed by +1 and −1 respectively, so the percentage of NRFU households with rosters differing in count by ±1 was 87%. The distributions for both Self-Response and for NRFU cases were relatively symmetric.

Other findings reported by Mulry et al. (2023a,b) include: (1) the earlier the data were acquired, the higher quality and therefore the greater the agreement between census and administrative records; (2) closeout NRFU cases were of lower quality than noncloseout NRFU cases; (3) the percentage of cases in which the census roster had a higher count than the administrative records roster varied as a function of the race or ethnicity of the householder; (4) the extra census person not found on the administrative records roster for NRFU households was spread among householders, spouses, relatives, children, unmarried partners, and nonrelatives; (5) there was a clear spike at age 17 for the age of the additional census person; (6) for one-visit NRFU households with multiple-source administrative records rosters in which the administrative records roster had one additional person in comparison with the census roster, the extra person was resident at the same address 98% of the time, though over 75% of the additional people were also found at other addresses, suggesting either a census undercount, movers, or people with attachments to multiple housing units; and (7) when the difference between the census count and the administrative records roster is 1, the census roster is larger than the administrative records roster between 55–95% of the time, depending on the race/ethnicity of the head of household, the race/ethnicity of each resident, whether the spouse, unmarried person or children are present, and the ages of the residents.[4]

When examining the nature of the additional person on either the administrative records roster or the census roster, the impact of the COVID-19 pandemic on increasing the percent of movers must be kept in mind. The pandemic also increased the number of people leaving colleges and universities earlier in the year, increased the number of evictions because of job losses

[4]Analyses of who the extra person was when administrative records rosters have an additional resident in comparison to the census roster are understandably much more limited given the lack of information on administrative records forms.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

and subsequent inability to make rent payments, and increased the number of people leaving large cities for less-populated areas, to avoid exposure to the virus.

8.5 WHAT THE EVALUATION DOES NOT ANSWER

Given the stated aims of this particular use of administrative records data in the 2020 Census—to reduce the costs of NRFU field work while not detracting from the quality of the counts generated—it is unfortunate that data are not available on cost reductions and that the nature of any resulting changes to the quality of the census counts remains unclear. It is unknown how the quality of the administrative records enumerations would have compared with the quality of the enumerations if those cases had been retained in NRFU. As Fay (2021) states: "the available evidence from the published record does not support a definitive answer to the question of whether use of AR [administrative records] enumerations reduced or improved the quality of the 2020 Census."

Limited information on the quality of administrative records enumerations is available from the 2020 Post-Enumeration Survey (PES) and from documentation of item nonresponse rates. This information is suggestive but just scratches the surface of what needs to be learned.

Based on the PES, in terms of correctness of enumeration, administrative records enumerations in 2020 were on a par with Self-Response and also with NRFU enumerations obtained from the household heads. The 2020 PES estimated that 97% of self-responses with an ID were correct, compared with 95% for administrative records enumerations, and 94% for NRFU enumerations with the household head. In contrast, only 87% of NRFU proxy enumerations and only 77% of NRFU enumerations with a household member other than the head were correct (see Table 3.4). Moreover, by definition, administrative records enumerations did not produce whole-household imputations, which occurred for every other enumeration type, even Self-Response.

In terms of item nonresponse, the picture was not as favorable. While only 1% of self-responses were missing age or date of birth and only 2–3% were missing race and ethnicity, 3% of administrative records enumerations were missing age or date of birth, 18% were missing race, and 28% were missing ethnicity. By comparison, NRFU enumerations with a household member had a higher percentage than administrative records enumerations missing age or date of birth (16%) but substantially lower percentages missing race (9%) and ethnicity (6%). Proxy NRFU enumerations were the worst—61% were missing age or date of birth, 41% were missing race, and 38% were missing ethnicity.[5]

[5]From 2020 Census Operational Quality Metrics Release 3, Table 2 (U.S. Census Bureau, 2021c). Item nonresponse rates refer to missing responses; item imputation rates are usually higher because they include cases in which responses had to be blanked and imputed. Denominators

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

The evaluation work done to date (Mulry et al., 2023a,b), while helpful, intriguing, and possibly the only type of assessment one could do regarding this component of the 2020 Census without making fairly heroic assumptions, does not greatly advance understanding of the causal nature of the differences between these two types of enumeration. Looking, for example, at administrative records and census rosters that differed by one (or more) people, it is unknown how many differing rosters were due to movers, students away at college, or people that had attachments to multiple households. It is also unknown how many differing rosters were due to the way administrative records rosters were formed, particularly those associated with personal income tax records. The main reason for this lack of understanding is that there is not a great deal of useful characteristics information attached to many of the administrative records files. NRFU was not pursued for even a sample of these administrative records-enumerated households during the 2020 Census, which further limits the possibility of obtaining a better understanding of the counterfactuals involved.

Questions in need of further investigation include the following:

1) How often did the administrative records model for housing unit status correctly identify nonexistent, vacant, and occupied dwellings? Even this question is not easy to answer since there are no true values for housing unit status (though the independent listing done for the PES could be used as a substitute for true values). Also, the quality of USPS UAA codes is unknown.
2) How does the quality of administrative records enumerations compare with the quality of self-responses? This does not bear immediately on how the use of administrative records enumerations affected the quality of the counts in 2020, but it could provide useful information about when to cut off Self-Response in the future.
3) How does the quality of administrative records enumerations compare to the enumerations collected during NRFU interview attempts 1–6, using several thresholds for the quality of administrative records?
4) How does the quality of administrative records enumerations compare to NRFU proxy enumerations? This could be answered for several thresholds for various demographic groups and types of housing (e.g., detached, multiunit, seasonal). The impact of needing corroboration could also be examined. Further, the possibility of using proxy information in conjunction with administrative records information for purposes of enumeration could be investigated. Finally, the possibility of using differing thresholds and differing degrees of corroboration depending on the urgency of closeout could be examined.

for percentages are people in occupied units providing the designated type of response (e.g., Self-Response).

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

5) How does the quality of administrative records enumerations compare to NRFU-generated whole-household imputations?
6) What are answers to questions 1–5 if conditioned on areas that are majority Hispanic people or majority non-Hispanic Black people? Fay (2021) mentioned that an examination of vacancy rates for block groups classified in this way led to a Census Bureau concern about the overestimation of administrative records-determined vacant housing units in those block groups. Fay (2021) noted, "Whether the quality of AR [administrative records] enumeration will vary in other respects for members of previously undercounted groups remains an important question." These questions could also be considered as a function of the threshold used to assess the quality of the administrative records used.
7) The training set for the occupancy models was 10 years old when applied, so are there alternative possibilities for a more recent training set?
8) How could the Census Bureau address the fact that administrative records enumerations in the 2020 Census were not completed in NRFU for interview attempts 2–6, so no training set based on 2020 data could provide information from comparisons of NRFU with administrative responses?

In addition, both the person-place model and the household composition model need to be validated using standard diagnostics, such as cross validation, to assess their degree of fit.

There is no question that the Census Bureau is to be applauded for its research program and first application of administrative records for enumeration, which likely resulted in improved census counts. However, questions do remain—the most important being the degree of variation in the quality of enumeration for race and ethnicity groups.

8.6 LOOKING TO 2030

8.6.1 Administrative Records Enumeration as Part of Nonresponse Followup

The use of administrative records for census enumerations in NRFU, while kept to a relatively modest 4.6% of total enumerations in 2020 (including vacant and nonresidential housing units), is potentially a key change for future census taking. There are good reasons to believe that this method would become more effective through increased use, especially given that the quality of administrative records may improve over time. The remaining questions are: (1) what share of enumerations in the 2030 Census could be completed using administrative records while still keeping with census data-quality standards; (2) what are the research studies needed to answer this question; and (3) what data

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

inequities are a consequence of the use of administrative records, if any, and how might they be reduced?

The complication in conducting such research is that there are no true values for the population counts for domains of interest. The ideal would be to compare the counts from a census using administrative records for enumeration to the same census *not* using this methodology, and then to compare *both* to the truth. But, without truth, one must be aware that both the administrative records entries and the census counts are subject to undercoverage, duplication, and erroneous enumeration, which complicates comparisons. This problem could be addressed by applying the results of the 2020 PES to produce a set of basic collection units (similar to block groups) for which there is reason to believe the counts are reasonably close to being correct. However, the time lag of the 2020 PES could potentially be a concern for this purpose.

The use of artificial (simulated) populations with a variety of properties—for example, various relations of administrative values to true values—is a valuable component of a comprehensive assessment of the performance of administrative records and should be pursued. The idea is to simply make a simulated census data set based on conjectured degrees of Self-Response, NRFU of various stages, MAF errors of various types, enumeration errors of various types, etc., and then test the performance of various models as the simulated dataset is adjusted for various assumed properties of the U.S. population and NRFU workload. The disadvantage is all the assumptions one must make, but the benefit is that one knows what the true values are for the simulation. If some robustness is evident across assumption sets, it could provide a degree of confidence that the findings are generalizable.

Given all the above, it seems that the best way of answering these questions is through a census test, in which a sample of cases that are enumerated using administrative records are carried through to the end of NRFU enumeration for purposes of comparison with a PES providing "true" values. Absent such a research effort, there will not be an evidence base for the Census Bureau to determine the extent to which this methodology should be used for the 2030 Census.

In terms of error assessments for such research, it is reasonable to treat the use of administrative records for enumeration as a type of whole-household imputation. In that case, to make a global comparison of the two techniques (as opposed to answering specific research questions as to why these two systems provide different counts), it is relatively unimportant to compare the number of occupants of individual households using census field processes to the number arrived at through estimation using administrative records—it makes more sense to compare the counts for small areas.

If one is interested in testing artificial populations, the application to this problem would be straightforward. Given the concern about the quality of administrative records-based enumerations for some demographic groups,

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

domains defined using such demographics could also be employed, and the resulting impact on errors examined.

While progress was made on the problem of incorporating administrative data into the enumeration of the 2020 Census for particular housing units for which this approach was supportable, there remain reasons to assume that the best approach for 2030 may differ in important respects from the approach taken in 2020. This is because there is no reason to believe that any of the current models—household unit status model, person-place model, and household composition model—or the distance function are optimal. For example, for the person-place model, one might replace the minimum with something less conservative. In the case of the distance function, some alternatives to consider are weighted Euclidean metrics, weighted absolute deviations, or probability of error with cutoffs. (See Efron, 1978, for additional ideas.) Also, it is unclear whether the optimal thresholds were applied. There are also questions about the appropriate census processes used in conjunction with administrative records data. For example: (1) Should every NRFU case receive a single NRFU visit prior use of administrative records data? (2) Should every initially determined vacant household be sent an additional postcard? (3) Should the Census Bureau separate various administrative records files into primary and corroborative groups? Given these and other questions, a research program is needed that assumes nothing about the best way to proceed. All these questions and other aspects of the use of administrative records in 2030 NRFU need to be examined and alternatives evaluated so the best version of the method can be applied.

Assuming continued research and evaluation can be done to establish the quality of administrative records-based enumerations of portions of the NRFU workload and to address other methodological and operational issues, it could well make sense for the Census Bureau to plan to expand the percentage of NRFU cases enumerated in this manner. A doubling from 5 to 10% of administrative records-based enumerations of addresses (which would be about 25% of the NRFU workload based on 2020 self-response rates) could be a challenging but feasible target. Of course, other applications of administrative records in the census, such as assistance in identifying new addresses, editing and imputation, improving the operational efficiency of data collection, and assistance in evaluation are important, not only to continue, but also to expand and improve to the extent possible. In addition, use of administrative records (e.g., birth records) to correct the persistently large net undercount of children ages 0–4 in the census (see Chapter 4) is an area to pursue.

Copyright National Academy of Sciences. All rights reserved.

8.6.2 Using Administrative Records on a Large Scale for Census Enumerations—Examples from Peer Countries

Peer countries that do not have well-developed population registers are conducting research on using administrative records to augment the traditional census (e.g., for imputation of missing items) and to enumerate a large portion—perhaps a majority—of the population. Surveys (self-responses) would be used for the remainder of the population in a "combined approach." Examples include:

- The United Kingdom (Henry et al., 2023) used administrative records for such purposes as imputation for missing items in its 2021 Census and is working toward a fuller use of administrative records for its annual population estimates for local areas with about 1,000–3,000 people.
- Canada (Bowlby and Beaulieu, 2023) is working to increase the percentage of administrative records enumerations from less than 0.1% of occupied housing units in 2021 to a significant percentage of the NRFU workload in 2026, and then to a census that combines self-response with administrative records in 2031. Alternatives being considered for 2031 include eliminating NRFU altogether or dividing the country into zones enumerated exclusively with administrative records and zones enumerated with a traditional census.
- Australia (Holmberg and Watmuff, 2023) used integrated person-level administrative records in its 2021 Census for such auxiliary purposes as improving imputation. It also had a backup plan (as did Canada) to use administrative records in areas that experienced wildfires or similar natural disasters at the time of the census. Australia is working toward the use of administrative records for enumeration on a small scale in 2026 and on a much larger scale for 2031.
- New Zealand (Bycroft and Matheson-Dunning, 2023) is working toward an "administrative records first" census with surveys for some information in 2028, driven by natural disasters (an earthquake delayed the 2011 Census to 2013 and cyclones have delayed the 2023 Census), unexpectedly low response rates in the 2018 Census, and high-cost field work in the 2023 Census. Its new Data and Statistics Act of 2022 gives authority to Statistics New Zealand to obtain records from public-sector agencies and drops language prescribing how census data were to be collected for both population and dwellings. New Zealand used administrative records in 2018 to create its dwelling frame, for item nonresponse, and for quality assessment. Because of low response, New Zealand ended up using administrative records for about 3% of the population in specific dwellings and another 8% in small geographic areas. By design, New Zealand is expanding the use of administrative records-based enumerations in 2023 and engaging in public discussions about

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

its preference to use a combined administrative-records-first-with-surveys model for 2028.

Bowlby and Beaulieu (2023) report the following from Statistics Canada's evaluations of its administrative records and models to date:

- Statistics Canada has access to good-quality administrative data but not for all areas or population groups;
- Administrative data population counts appear to overcount the population somewhat but less so than the traditional census undercounts the population;
- The current "in-scope" model accurately predicts if an individual is in or out of scope for the census 88% of the time;
- The current "household model" places individuals at correct addresses 80% of the time; and
- The large majority of census households can be exactly recreated by administrative data—61% of derived households match exactly to the census.

These assessments by Statistics Canada are encouraging but also indicate how far Canada still has to go to use administrative records enumerations with confidence in the quality of the resulting data. The United States would need similar fine-grained assessments for geographic areas and population groups.

8.6.3 Determining the Appropriateness and Feasibility of a Significantly Increased Role for Administrative Records Enumerations in the 2030 Census

Congruent with peer countries, the Census Bureau and groups contracted to consider new methods for the 2030 Census have signaled interest in the potential of using administrative records enumerations for a substantial portion of the population with follow-up surveys to fill in gaps (e.g., Macagnone, 2022; JASON, 2016 and associated news accounts such as Mervis, 2017; and Keller and Shipp, 2021). We do not have knowledge of the Census Bureau's specific research plans for possibly increasing the use of administrative records enumerations in a substantial way in 2030 and thus cannot comment on them, nor do we take a position on the general principle of original enumeration via administrative records. That said, as we discuss again in Chapter 12, we believe that such a major shift is almost certainly premature for the 2030 Census. In this section, we note the following kinds of technical, operational, political, legal, and ethical issues that need to be considered—issues that statistical offices in the UK, Canada, Australia, and New Zealand have identified as critical in their own research on administrative records for enumeration purposes, particularly the need for public and stakeholder outreach and engagement.

Copyright National Academy of Sciences. All rights reserved.

Data Quality

A key issue for assessing the potential for administrative records-based enumerations is the quality of currently and potentially available records, assessed in terms of such attributes as: timeliness (referencing Census Day and available to the Census Bureau in the needed timeframe); accuracy of identifying a living person at the person's usual residence according to census residence rules; and completeness and accuracy of characteristics. For equity in allocation of votes and resources, quality needs to be assessed not only overall but also for geographic areas and population groups. Thus, for legislative redistricting, the United States has statutory and case law requirements for accurate data for small areas and race and ethnic populations. Quality also needs to be assessed both in absolute terms, relative to a measure of truth, and in comparison with other types of enumerations, such as NRFU household and proxy interviews and whole-person imputations.

Public and Stakeholder Acceptance

In the United States, the census is required by the Constitution, with Congress having authority over the content and methods. Title 13 delegates this responsibility to the U.S. Secretary of Commerce, but Congress can and has exerted its will—to add race categories, for example, and to mandate that the Census Bureau provide small-area data to states for redistricting. According to Title 13, the census of population is also a census of housing (first conducted in 1940). Regarding items in the census, a review of federal agency requirements determined that all items, including age, sex, race, ethnicity, household relationship, and housing tenure, were mandated in legislation for subnational geographic areas, such as census tracts and school districts (National Research Council, 1995:App. M).

Regarding administrative records use, Title 13, Section 6 provides authority for the Secretary of Commerce to obtain records of federal agencies and other entities for use in its censuses and surveys "consistent with the kind, timeliness, quality and scope of the statistics required." Nonetheless, given the long history of "actual enumeration" for the census, and of adverse reactions to such concepts as data warehouses (see, e.g., Yorsz, 1976), an active and early dialogue with Congress would be essential to achieve buy-in for any Census Bureau plan involving substantial use of administrative records for original census enumeration. Indeed, legislation allowing such use could be desirable to avoid litigation over the meaning of "actual enumeration."

Dialogue would also be needed with other stakeholders, including federal agencies, state and local governments, and data users in government, academia, the private sector, and the nongovernmental organization sector. Civil rights groups and others would be particularly concerned with how well

Copyright National Academy of Sciences. All rights reserved.

administrative records cover areas and population groups and what means would be used to fill in gaps. Finally, the general public would need to be consulted regarding perceptions of a combined census approach as opposed to the concept of the census as a civic ceremony in which personal participation by all is important and expected.

A related issue that would need stakeholder and public input concerns the move from a household-based census to a person-based census, which use of administrative records entails. There are many important user needs for data not only on housing units and people, but also on households and families. For example, knowledge of household and family structure in terms of numbers of children and adults is critical for many services, including education. Historically, the census has been a nested data collection, using the address as the basic enumeration unit and determining the occupancy status of housing units, and the people resident in occupied units and their relationships. Administrative records may contain only individual-level data (e.g., Medicare) or data for disparate program units (e.g., income tax filing units or income support program eligible units). The Census Bureau's models described above link people to housing units and addresses, but their quality needs to be assessed.

One more issue concerns the choice of administrative records. Brown et al. (2023) used a large number of administrative records to simulate a "real-time" 100-percent-administrative-records census (see Section 8.6.4). The records included those authorized under a 2019 Executive Order intended for the Census Bureau to estimate the citizen population, such as U.S. Department of Homeland Security records. For use as census enumerations, it would be essential to take ethical considerations into account in selecting records so that the trust of the public and stakeholders is not undermined by the use of records perceived with distrust and apprehension.

Operational and Technical Considerations

A "combined approach," which uses administrative records enumerations and surveys to fill in gaps, appears to be the goal of countries such as Canada and Australia. A challenge for a combined approach is how to handle messaging, self-response targeting, and field operations for areas and groups that are not well counted by administrative records and thus require surveys. It is easy to imagine considerable public confusion with likely adverse effects for cooperation. Another possible operational problem involves the potential for an important administrative records source to become unavailable or substantially altered close to the census.

It follows that, to generate confidence in the use of administrative records data for original enumeration of households, it is even more imperative to resolve technical concerns with the models associated with using administrative records to supplement NRFU. As pointed out in this chapter, research is needed

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

on (among other things) bridging the long 10-year gap between U.S. censuses in using the previous census as the training set for inference. We also noted that, logically, the 2020 Census will be unable to serve as a training set for original enumeration via records.

8.6.4 "Real-Time" Simulation of an Administrative Records Census

The Census Bureau recently released *Real-Time 2020 Administrative Record Census Simulation: A Design for the 21st Century* (Brown et al., 2023), one of the planned evaluations in the formal program of evaluations and experiments for the 2020 Census. The continuation of work that has been conducted for over 25 years at the Census Bureau on the application of administrative records for various uses in censuses and surveys (and been an experiment in 2000 and 2010 Censuses), Brown et al. (2023) considers the problem of producing a person-based population count from administrative records in place of a traditional census. It is crucial to differentiate, again, this person-based approach from the methodology actually used in the 2020 Census: using administrative records data in a small percentage of NRFU cases for which at least one in-person NRFU attempt was made and for which the quality of administrative records information was judged to be high. By comparison, Brown et al. (2023) did not attempt to enumerate the occupants of residences but instead computed a sum of weighted averages of probabilities of occupation over possibly several residences.

The administrative records census simulated by Brown et al. (2023) exceeded the count achieved in the 2020 Census by 2.3%, and it is certainly a useful analytic exercise to guide further research and development. However, in several respects, it falls short of the quality standards that would be needed for actual implementation in a decennial census. First, the method used in the simulation added into the census people for whom the location of residence was unverified or, if verified, not geocoded. As Brown et al. (2023:xiii) pointed out: "About 14.4 million administrative record people are observed at 9.8 million addresses outside the 2020 Census collection universe, and another 8.7 million are at addresses not linked to the MAF." Where would such households be placed geographically for reapportionment and redistricting? Also, as mentioned above, the algorithm used often did not place individuals in single households. Instead, it used a set of weights for each person based on probabilities that each individual was a resident of various households. The resulting uncertainty, especially for low levels of geographic aggregation, can be seen in Brown et al. (2023:Table 69) for aggregates at the tract level and below. This uncertainty is inconsistent with data users' needs for small-area data from the census.

Second, the characteristics information necessary for various applications, including redistricting, often had to be modeled or imputed. It is difficult to compare the degree of modeling or imputation carried out in Brown

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

et al. (2023) for characteristics information with the degree of missingness of characteristics information in the 2020 Census responses. This is because in Brown et al. (2023), a characteristics variable could be missing on all of the administrative files used, or multiple disagreeing values could be available for a characteristic. While comparisons are complicated, the degree to which unique characteristics were unavailable in the group of files was a quality issue that Brown et al. (2023:248) acknowledged: "More research is needed to improve business rules and models for predicting demographic characteristics of people in administrative records. The rules and demographic modeling for administrative records in this report do not match well with as-reported values in the 2020 Census." Further, Brown et al. (2023:Table 77) showed that characteristics imputation alone was responsible for a considerable amount of uncertainty for race and ethnicity values.

Third, Brown et al. (2023:247) conceded that matching error, both false nonmatches and false matches, are a concern using this method. They recommended that:

> Record linkage error should be quantified and minimized. AR [administrative records] census estimates contain record linkage error. Though we believe the remaining duplication in the AR census after unduplicating person records by PIK/EPIK[6] is likely to be very low, the magnitude of the false nonmatches (as well as false matches) should be estimated. It would be helpful to conduct research on the duplication rate among PIKs and EPIKs and on how entity resolution could be improved to decrease that rate. It would also be desirable to develop record linkage error measures for the Census Bureau's production record linkage system and to estimate how the errors propagate in the statistics.

In addition, as noted previously, some of the administrative record files used in the study could create considerable fear and distrust if used in the census. For instance, the 31 administrative lists used included: (1) Immigration and Customs Enforcement (ICE)-provided data from its Enforcement and Removal Operations; (2) ICE-provided data from the Student Exchange Visitor Information System; (3) U.S. Citizenship and Immigration Services (USCIS)-provided data on non-U.S. citizens receiving Deferred Action for Childhood Arrivals, those classified as Special Immigrant Juveniles, and those otherwise interacting with USCIS and appearing to lack a lawful immigration status; (4) USCIS-provided data on non-U.S. citizens with Temporary Protected Status;

[6]Enhanced Protected Identification Keys" (EPIKs) are unique identifiers attached to persons, and thereby represent an extension of PIKs, through the use of additional reference files added to the Person Validation Identification System (PVS). EPIKs enable the Census Bureau to provide unique codes for residents of the United States who, without such an extension, would be unable to be provided with such identification codes. Doing this greatly facilitates record linkage for these individuals across various administrative records, which is used to identify their duplicates across administrative records systems.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

and (5) U.S. Customs and Border Protection-provided data from the Arrival and Departure Information System. It is easy to see how use of some of these lists could seriously impact public opinion concerning the census. Brown et al. (2023) noted that these types of records (i.e., from ICE) were available to the Census Bureau pursuant to Executive Order No. 13880 (2019) that was subsequently repealed and that these records have not been approved for future use.

Finally, and perhaps most importantly, there is currently no comprehensive assessment of the quality of an administrative records-based approach in comparison to that of the current methodology of the decennial census. While the total count of 2.3% above the 2020 Census is encouraging, there are larger discrepancies for smaller geographic areas, and there are sizeable discrepancies for various demographic groups. Beyond what is provided in the Brown et al. (2023) study, fuller understanding is needed of the differences between the two methods as a function of the various types of households and housing units, as well as the impact such an approach would have on public opinion and on the uses of these estimates for standard decennial census purposes.

8.7 CONCLUSIONS AND RECOMMENDATIONS

Conclusion 8.1: Based on the information currently available, the U.S. Census Bureau deserves commendation for the meaningful steps made in the limited use of administrative records as an alternative to the enumeration of some nonresponding households in the 2020 Census. The Census Bureau acted with appropriate levels of both boldness and caution in settling on conditions under which a nonresponding household at a designated residence would be enumerated using records information after only one enumerator visit rather than six, yet still ensuring that single in-person visit. Moreover, this change in approach was preceded by an extensive history of high-quality experimentation and testing.

Recommendation 8.1: The U.S. Census Bureau should continue its research and development program concerning the best ways to use administrative records as a supplement to decennial census operations. Potential uses of administrative records include expanding enumeration of limited subsets of the 2030 Census population in the Nonresponse Followup workload, reducing proxy responses and whole-person imputations, and possibly redressing the long-standing net undercount of children ages 0–4.

Conclusion 8.2: Moving to the use of administrative records to enumerate a substantial proportion of the population without a contact

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

attempt does not seem feasible for the 2030 Census. Rather, it appears to be a long-term proposition, requiring extensive testing, research, and engagement with Congress, other stakeholders, and the public. Not only are there operational and technical issues to resolve and data-quality standards to determine and assess for geographic areas and population subgroups, but there is also the need to consider the ramifications of switching from a household/address-based focus to a person-based focus and to address issues of legal, constitutional, and public acceptability. It will be essential for the U.S. Census Bureau to be as transparent as possible about its research and testing and to have open and constructive dialogue with all parties.

Copyright National Academy of Sciences. All rights reserved.

– 9 –

Measurement of the Group Quarters Population

9.1 WHAT ARE GROUP QUARTERS AND WHY ARE THEY IMPORTANT?

Group quarters (GQs) are places where people live or stay in a group living arrangement, and that are owned or managed by entities or organizations providing housing and/or services for the residents. These services may include custodial or medical care as well as other types of assistance. GQs differ from typical household living arrangements because the people living in GQs are usually not related to one another. GQs include correctional facilities, college/university student housing (e.g., dormitories and Greek society houses), and nursing homes and other health care facilities, as well as a variety of other places both institutional (e.g., psychiatric hospitals) and noninstitutional (e.g., group homes and residential treatment facilities) in nature. Moreover, GQs also include locations where people who have no fixed addresses receive services, such as shelters for those experiencing homelessness.

With just 8.2 million persons in a nation of more than 333 million, at first glance it seems reasonable to conclude that any analysis of GQs would be inconsequential for an evaluation of 2020 Census quality. That is hardly the case, however, for two reasons. The first has to do with the outsized impact of the GQ population, due to the very uneven distribution of GQs throughout the country. Second, GQs are not a monolith but a combination of various groups of people that are very different in their living arrangements and other characteristics, as with those living in college/university housing, correctional

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

facilities, and nursing homes. Thus, each GQ type usually requires a unique set of procedures, rules, and processes for census data collection. Further, when GQ populations are enumerated incorrectly, the impact can be severe—not only on population totals, but on the distribution of persons by characteristics. Consider the influence of a skilled nursing facility or a college dormitory on the age distribution of a small town, for example.

9.2 ACCURATE COUNTS OF GROUP QUARTERS POPULATIONS MATTER FOR LOCAL GOVERNMENTS

Although persons in GQs are a small percentage of the total population—just 2.5–3.0% for the last four censuses—there is wide variation by locality. The GQ population may be negligible in some places, while in others it is the dominant population, as when localities host a college, correctional facility, or other large GQ facility. For these localities' own uses of census data, and for uses that involve shares of a fixed pie (e.g., redistricting, fund allocation), accurate listing of GQ locations and accurate enumeration of GQ residents are critical census functions.

It is instructive to illustrate the local importance of GQs using two examples from the 2020 Census P.L. 94-171 redistricting data file, issued in August 2021:

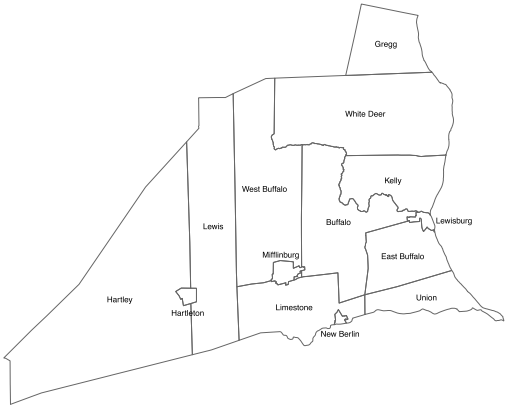
- *College towns*—In the 50 incorporated places with the highest proportion of students in college housing, students accounted for 35–95% of the population. Some examples are: Shepherdstown, WV, home to Shepherd University, which accounted for 47% of about 1,500 people; Prairie View, TX, home to Prairie View A&M University, which accounted for 71% of about 8,200 people; and Alfred Village, NY, home to Alfred University and Alfred State College, which accounted for 78% of about 4,000 people.
- *Correctional facilities towns*—In the 50 incorporated places with the highest proportion of people in adult correctional facilities, the residents accounted for 45–88% of the population. Some examples are: Florence, AZ, site of several federal, state, and private prisons, where incarcerated persons accounted for 57% of about 26,800 people; Sheridan Village, IL, site of Sheridan Correctional Center, where incarcerated persons accounted for 68% of about 1,700 people; and Nicholls City, GA, site of the Coffee County Correctional Facility, where incarcerated persons accounted for 76% of about 3,100 people.

The presence of a large GQ in a small governmental jurisdiction (e.g., place or minor civil division) can result in age, race/ethnicity, and sex profiles that are strikingly different from those of similarly small jurisdictions without such facilities. Such differences underscore the importance of correctly locating and enumerating GQ populations; see Box 9.1 for an example in Pennsylvania.

Copyright National Academy of Sciences. All rights reserved.

Box 9.1 Group Quarters in Pennsylvania Townships and Counties

In 2020, group quarters (GQ) residents were 3.1% of Pennsylvania's total population of 13 million (compared with 2.5% for the nation). This average obscures a highly skewed distribution at the local level. Of Pennsylvania's 2,572 minor civil divisions (5,000 average population, 1,800 median population), over one-half had no GQ residents in 2020. Of the 1,184 geographies with at least one GQ resident, fully 75% had fewer than 3.3% of their population in GQs. At the other extreme, in 10% of these geographic areas, the GQ population was between 9.7–21.6% of all persons, while 5% had between 21.9–83.3% of their population in GQs.



Diving deeper, four adjacent townships at the eastern edge of Union County, Pennsylvania, differ dramatically in GQ populations. East Buffalo township and Lewisburg borough both include student housing for Bucknell University—in 2020, Bucknell students in college housing made up 19 and 37% of the populations of East Buffalo (7,400 total persons) and Lewisburg (5,200 total persons), respectively. The township next door, Kelly township (4,000 persons total), houses the U.S. Penitentiary, Lewisburg, for men, which accounted for 15% of its population. The next adjacent township, White Deer (4,000 people total), had only 0.5% of its population in GQs in 2020.

9.3 NUMBER OF PERSONS IN GROUP QUARTERS BY TYPE

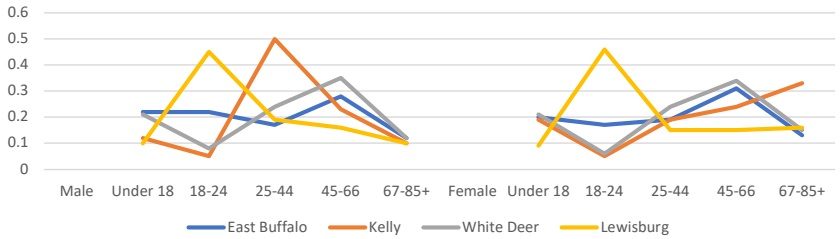
It is impossible to capture the diversity of GQ populations in a simple set of categories. Nonetheless, Table 9.1 divides GQs into seven types used by the U.S. Census Bureau, with each major category encompassing several subcategories,[1] many of which have unique populations requiring specialized data-collection methods. Three major types account for more than 77% of all persons in GQs: correctional facilities for adults, college/university student housing, and nursing facilities.

[1]For a detailed list of all subcategories, see Appendix Table A-1.

Copyright National Academy of Sciences. All rights reserved.

Box 9.1 (continued)

The figure below shows dramatically different age-sex distributions for these four geographies based on 2010 Census results (as of this writing, 2020 Census data were not yet available by sex and detailed age). Were there to occur a geocoding or other type of error that affected the location of Bucknell University housing or the Lewisburg Penitentiary, the consequences for the total population and age-sex distributions among these jurisdictions could be significantly affected.



SOURCE: Data obtained from the 2020 Census P.L. 94-171 Redistricting File, Tables P1 (total population) and P4 (group quarters population by type of group quarters); and the 2010 Summary File 1, Tables P12 (total population by sex and detailed age) and P43 (group quarters population by type, sex, and broad age category).

- *College/university student housing:* Students or family members of any age who reside in living quarters affiliated with a college or university comprise the largest group, with almost 2.79 million persons or about one-third of the GQ population in 2020.
- *Correctional facilities for adults:* Those who are under correctional supervision of some type (i.e., adjudicated as opposed to pretrial in a local jail) constituted the next-largest GQ group, 1.97 million or 24% of the GQ population.
- *Nursing facilities:* Those in nursing homes, including those requiring skilled nursing care, numbered 1.63 million persons, making up about 20% of all persons in GQs.
- *Other noninstitutional GQs:* More than 1.36 million persons fall into this category, a little less than 17% of all persons in GQs. The largest subcategories include group homes and residential treatment facilities for adults, as well as those without a permanent residence, such as persons who receive services in shelters, soup kitchens, and from mobile food vans, and who are counted in the Service-Based Enumeration (SBE) operation. Also included here are those enumerated in targeted non-sheltered outdoor locations (TNSOLs).

Copyright National Academy of Sciences. All rights reserved.

Table 9.1 Change in Group Quarters Population by Major Group Quarters Type, 2010 and 2020 Censuses, United States

	Population		Change, 2010 to 2020	
Group Quarters Type	2010	2020	Number	Percent
Institutionalized population	3,993,659	3,753,401	−240, 258	−6.0
Correctional facilities for adults	2,263,602	1,967,297	−296, 305	−13.1
Juvenile facilities.	151,315	88,115	−63, 200	−41.8
Nursing facilities/skilled-nursing facilities	1,502,264	1,627,046	124, 782	8.3
Other institutional facilities	76.478	70,943	−5, 535	−7.2
Noninstitutionalized population	3,993,664	4,485,615	491, 951	12.3
College/university student housing	2,521,090	2,792,097	271, 007	10.7
Military quarters	338,191	328,372	−9, 819	−2.9
Other noninstitutional facilities	1,134,383	1,365,146	230, 763	20.3
Total	7,987,323	8,239,016	251, 693	3.2

NOTES: Unless indicated otherwise, tabulations in this analysis are for the 50 states and the District of Columbia. Totals may differ slightly between tables in this chapter because of the application of noise and rounding to maintain the confidentiality of respondents.

SOURCE: U.S. Census Bureau tabulations from 2010 Census Advance Group Quarters Summary File (updated) and 2020 Census Redistricting Data (P.L. 94-171) Summary File, published at https://www.census.gov/library/stories/2021/08/united-states-group-quarters-in-2020-census.html.

- *Military quarters:* This category includes active-duty military living in military barracks located on a base or aboard a ship. They numbered about 328,000 persons or 4% of all persons in GQs.
- *Juvenile facilities:* Group homes for juveniles (noncorrectional), which account for about 1% of persons in GQs, include community-based group-living arrangements for youth in residential settings; residential treatment centers for juveniles (noncorrectional) for the treatment of drug/alcohol abuse, mental illness, and emotional/behavioral disorders; and correctional facilities intended for juveniles, such as facilities that provide strict confinement for their residents and detain juveniles awaiting adjudication, commitment or placement, and/or those being held for diagnosis or classification.
- *Other institutional GQs:* Adults who require assistance with activities of daily living and have difficulty functioning independently are just under 1% of all persons in GQs. These GQs include: psychiatric hospitals and psychiatric units in other hospitals; hospitals with patients who have no usual home elsewhere; in-patient hospice facilities; active-duty military in hospital; and live-in residential schools for persons with disabilities.

From 2010 to 2020, the total GQ population increased by 251,693 persons (a 3.2% change), with change unevenly distributed among the seven primary

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

types of GQs (Table 9.1). For the three largest GQ categories there were sizable changes: persons in correctional facilities for adults declined by 13%, while increases occurred for those in nursing facilities (8.3%) and in college/university student housing (11%). Those in other noninstitutional facilities increased by 230,000 (20%) between 2010 and 2020, a reflection of increases in the number of persons enumerated in small group homes, residential treatment facilities, and in the SBE.

The largest percent change between 2010 and 2020 occurred among those in juvenile facilities—a decline of 41% or 63,000 persons. A definitional change occurred between 2010 and 2020 that affects this comparison—those juveniles in residential treatment centers for mental health conditions were counted at their home residence in 2020, rather than at the facilities where these individuals were counted in 2010 (Fontenot, 2018b).

9.4 GROUP QUARTERS ADDRESS LIST DEVELOPMENT

In the development of the 2010 Census, the process of building the initial address list for Group Quarters Enumeration (GQE) was very field intensive. Living quarters were roughly classified as housing units, GQs, or transitory locations (TLs) during the all-field Address Canvassing operation, and a separate Group Quarters Validation (GQV) operation sent field staff to visit individual GQ locations, to more precisely determine the GQ type and obtain initial information about the size of the GQ. Consistent with the desire to make address updating more efficient, the Census Bureau opted not to reinstate GQV but instead defined a GQ/TL suboperation within In-Office Address Canvassing (IOAC), with the intent of validating GQ addresses in areas of change as identified by the Active Block Resolution (ABR) operation. Much of the work would be done in-office, using administrative data, local geographic information systems files, and commercial information, coupled with phone calls to GQ facility contacts (Fontenot, 2021g; Richmond and Hanks, 2022).

However, as described in Chapter 2, the ABR component of IOAC was discontinued in early 2017 due to funding uncertainty—with major implications for the GQ/TL list-development operation. No longer able to build on ABR work, the GQ/TL component was redesigned to try to review all GQ/TL entries on the MAF—albeit purely as a telephone-based operation using a standard calling script. As documented by Richmond and Hanks (2022), the GQ/TL list-development operation encountered problems from its start in summer 2017—much of the difficulty attributable to GQ facility respondents' skepticism over the validity of the operation given the early date (the 2020 Census not being high in public view in 2017). Not least of the problems was a subtle, cosmetic one—the fact that the dialing system was such that respondents' Caller ID systems failed to identify the call as being from "U.S. Census Bureau."

Copyright National Academy of Sciences. All rights reserved.

The problems were sufficiently major that the GQ/TL component of IOAC was itself terminated early, in March 2018, citing continuing funding concerns. Moreover, there was sufficient concern over the limited information that had been collected thus far—in particular, the suspicion that the small set of results already gathered was laden with duplicate entries—and the Census Bureau opted not to forward the IOAC GQ/TL results to update the Master Address File (MAF) (Richmond and Hanks, 2022:72).

Difficulties continued in In-Field Address Canvassing (IFAC) as well, due principally to field lister error or, at least, unfamiliarity with the concepts at hand. The Census Bureau's assessment of IFAC (Lane et al., 2022) notes that IFAC listers routinely inflated the counts of both GQs and TLs by counting each individual unit (i.e., every individual room in a college dormitory) as a GQ when only the parent site/facility should have been labeled a GQ. Moreover, quality-control work suggested that unit-type conversions by IFAC field staff (e.g., housing unit to GQ, or vice versa) tended to be problematic. Accordingly, the Census Bureau opted for the same solution as in IOAC, which was to largely refrain from using IFAC GQ/TL results to directly update the MAF—paring back clusters of newly added GQs to a single entry and treating unit-type conversions as verified housing units to be resolved in further field work (Lane et al., 2022:91).

Part of the rationale for not using 2019's Address Canvassing information to directly update GQ information on the MAF is that there remained a key opportunity for verification and review of the GQ list prior to census operations: the Group Quarters Advance Contact (GQAC) operation, a carryover from the 2010 Census design. In GQAC in January–February 2020, field staff contacted each GQ, drawing from GQ records already on the MAF and updates from state and local governments (i.e., the Local Update of Census Addresses program).[2] The GQAC interview verified the GQ name and type, collected address and contact information, and elicited an expected population count for April 1, 2020. Importantly, it also served to broker the eventual enumeration, asking which enumeration option (e.g., in-person interviewing, electronic response) would be preferred and attempting to schedule an appointment to perform the actual enumeration. Parallel work occurred for the enumeration of persons in TNSOLs, with an advance contact operation run between February 6–21, 2020, commencing final planning for the count.

[2]Experimentation with contact by phone to verify the GQ list was conducted as part of the 2018 Census Test in Providence, Rhode Island. In that test, the Census Bureau found that phone contact was not effective and, as a result, decided to adopt in-person visits for the 2020 Census GQAC verification operation. The GQAC operation for the 2018 Census Test was conducted between June 18–July 10, 2018, for an enumeration in July and August of the same year. See https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/plan/testing-activities/2018-census-test/test/providence/group-quarters/gqac.html

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

This advance work fell to the wayside in March 2020 with the arrival of the COVID-19 pandemic, the shutdown of all census field operations, and the major subsequent disruption to GQ living situations—not least of which the abrupt emptying of college and university campuses and sudden sensitivity to group gatherings of all sorts. As the Census Bureau resumed field work, it had to readminister the advance contact operations for GQs, TNSOLs, and TLs to account for changed situations. In the GQAC context in particular, as we describe later, the new contacts involved a concerted push to get GQs to use electronic (and field-staff contactless) response options as much as possible. In the case of TNSOLs, another round of advance contact was held August 17–September 11, 2020, with the field enumeration beginning on September 22.

9.5 NATURE AND ORIGIN OF GROUP QUARTERS ADDRESSES

In the analysis of GQ addresses in the MAF Development Dataset (examined intensively in Chapter 5), the wide range of complexity and nuance associated with GQ addresses became apparent. Certain GQ addresses remain stable over long periods of time, such as for large correctional facilities or longstanding college/university student housing or nursing facilities, while others are far more malleable, such as small group homes for adults, residential treatment facilities, and facilities that provide services to those without a permanent address, such as shelters for people experiencing homelessness. As a consequence, no single *precensus* program can be expected to capture all addresses needed to conduct a complete enumeration of the GQ population.

Table 9.2 shows the origins of GQ addresses in the 2020 Census. Close to 56% were from previous censuses and older Delivery Sequence Files from the U.S. Postal Service. The largest single category of addresses resulted from GQ operations themselves (GQAC and GQE): 93,500 MAF units or 35% of the total, with another 7% from Local Update of Census Addresses submissions, Address Canvassing[3] and Update Leave/Transitory Location operations. While many GQs are older and established, others are "moving targets" for data collection, with the latter requiring concerted listing efforts around the time of or even during the census.

Thus, there is a "churn" of GQ addresses around census time, which is largely the product of the more malleable group, as when the Census Bureau pays special attention to areas with large numbers of new housing units. Table 9.3 shows the distribution of the more than 95,000 MAF units involved in "transitions" during the census, in which the final disposition was a GQ MAF unit (e.g., a housing unit that transitions to a GQ during the enumeration), or in which a GQ was added as a result of one or more census address-building

[3]Despite rejecting the results of Address Canvassing for the validation of GQs, a small number of those addresses ultimately made it into the 2020 GQ address list.

Copyright National Academy of Sciences. All rights reserved.

Table 9.2 Origins of Group Quarters Addresses in the 2020 Census

Original Source of MAF Address	Number	Percent
On MAF Before 2020		
2000 Census and Earlier	67,000	25.3
2010 Census	46,000	17.4
Pre-2020 Delivery Sequence Files	34,500	13.0
Other Pre-2020 Operations & Tests	5,200	2.0
2020 Census Operations		
Group Quarters Operations	93,500	35.3
LUCA Program	11,000	4.2
Address Canvassing	6,600	2.5
Update Leave/Transitory Locations	1,000	0.4
*Total Group Quarters MAF Units**	265,000	100

* Estimates are based on MAF Units from the MAF Development Dataset and differ slightly from the number enumerated in the census.

NOTES: MAF, Master Address File; LUCA, Local Update of Census Addresses

SOURCE: 2020 Census MAF Development Dataset. See Disclosure Review Statement; CBDRB-FY23-0221.

operations. More than 94% of this "churn" is dominated by two categories, both involving the identification of new GQ MAF units—that is, units that did not exist in the MAF prior to enumeration activities and were added, most as a result of the GQAC operation.

The large majority of GQs in these two largest categories are considered "other noninstitutional" GQs, which mostly consist of group homes for adults and SBE sites, especially those classified as TNSOLs.[4] These are the true "moving targets," both from the standpoint of listing and enumeration. Their changing nature makes them difficult to detect ahead of time and to classify correctly, which means that good operations will always be needed in the years immediately prior to—and even during—the census to cover them well. Underscoring the difficulty, a sizable number of these MAF units, especially TNSOLs, were recoded as having 0 population in the 2020 Census.

9.6 POSTCENSUS REVIEW

Before delving into the mechanics and content of GQE in the 2020 Census, it is useful to continue the thread of GQ list/frame development, skipping ahead briefly to the aftermath of census data collection and discussing postcensus GQ review.

[4]Email correspondence with Stuart Irby of the U.S. Census Bureau Geography Division, November 23, 2022.

Copyright National Academy of Sciences. All rights reserved.

Table 9.3 Transitions and Additions for Group Quarters Master Address File Units in the 2020 Census

Transition or Addition	MAF Units	Percent
Sent to enumeration as HU, classified as vacant GQ	1,200	1.3
Sent to enumeration as HU, enumerated as GQ	2,200	2.3
Sent to enumeration as TL, classified as vacant GQ	100	0.1
Sent to enumeration as TL, enumerated as GQ	200	0.2
Added (existing MAF unit) by census and classified as vacant GQ	50	0.1
Added (existing MAF unit) by census and enumerated as GQ	1,600	1.7
Added (new MAF unit) by census and classified as vacant GQ	26,000	27.2
Added (new MAF unit) by census and enumerated as GQ	64,000	67.0
	95,500	100.0

NOTE: MAF, Master Address File; HU, housing unit; GQ, group quarters; TL, transitory location. Table covers MAF units with a final disposition as a GQ location.

SOURCE: Special tabulation from the MAF Development Dataset by Stuart Irby, Geography Division, U.S. Census Bureau. See Disclosure Review Statement; CBDRB-FY23-0221.

In the 2010 Census, the Census Bureau—working with the Federal-State Cooperative on Population Estimates—conducted a check of GQ addresses to identify missing GQs, holding this review late in the data-collection and processing stages so that problems could be fixed *before* the release of apportionment counts on December 31, 2010. The Census Bureau hosted a group of local officials who had Special Sworn Status and access to local data, to review interim GQ data and assess deficiencies.

After the successful implementation in 2010, a similar late-stage check of GQ data was planned in the 2020 cycle as adjunct to the Count Review program. However, the COVID-19 pandemic-induced delays and tight operational deadlines—as well as restrictions on in-person gatherings—caused the GQ review program to be conducted on a much more limited scale. Ultimately, the release of the apportionment counts was delayed beyond December 31, 2020, but it was too late to perform the GQ review program. In lieu thereof, and acknowledging likely problems with the count, the Census Bureau worked with local stakeholders to construct a post-census review program called Post-Census Group Quarters Review (PCGQR). While this review program could not prevent GQ coverage problems entering the census tabulations, it provided an opportunity for local governments to identify missing GQ addresses and for the Census Bureau to correct the base for population estimates going forward. See Box 9.2 for a description and current status of the program.

This experience has major ramifications for any program aimed at preparing for the 2030 GQ enumeration. While it is important to curate a list of GQs throughout the decade with some equivalent of an ongoing GQAC

Copyright National Academy of Sciences. All rights reserved.

Box 9.2 Post-Census Group Quarters Review Program

The 2020 Post-Census Group Quarters Review (PCGQR) program provided a mechanism for tribal, state, and local governments, or their designated representatives, to request a review by the U.S. Census Bureau of group quarters (GQ) populations that were believed to be counted incorrectly in the 2020 Census. The PCGQR program started June 6, 2022, and ran through June 30, 2023. Governments could submit records showing GQ populations for facilities missed entirely in the 2020 Census or facilities for which populations were counted incorrectly as of April 1, 2020. All GQs were in scope, except those in Service-Based Enumeration locations such as soup kitchens, mobile food van locations, and targeted non-sheltered outdoor locations (TNSOLs). If the review process found discrepancies in these population counts that were supported by sufficient documentation, approved revisions to the GQ population were provided to the Population Estimates Program for incorporation into the base population for upcoming vintages of estimates. PCGQR had no effect on the numbers for 2020 Census products.

As of July 6, 2023, 69 jurisdictions from at least 22 states made PCGQR submissions to the Census Bureau, including the cities of Boston, Chicago, Detroit, New Orleans, New York, and Phoenix, and the states of California, Illinois, and Texas; see https://www2.census.gov/programs-surveys/decennial/2020/program-management/pcgqr/pcgqr-cases-received.pdf.

program, it is unrealistic to expect that such an activity alone will provide good coverage of the GQ universe for 2030. A vital part of good coverage needs to involve verification and discovery activities immediately prior to and during the enumeration itself.

9.7 THE COLLECTION OF GROUP QUARTERS DATA IN THE 2020 CENSUS

9.7.1 Occupancy Status

The changing nature of GQs in and around the 2020 Census enumeration is reflected in their occupancy status, which is shown in Table 9.4. The most salient finding was the relatively high level of vacancy overall, with 23% of GQs deemed vacant during the census, and another 4% subject to deletion. Except for nursing facilities, vacancy levels were in the double digits for all major GQ types. It is important to keep in mind that vacant GQ units are part of the census enumeration, not unlike vacant housing units with recognized addresses that are part of the housing stock, irrespective of occupancy status.

Owing to the unpredictable nature of military deployments and changing status of military bases, 39% of military GQs were deemed vacant in the enumeration. In contrast, nursing homes and college/university student housing had lower levels of vacancy, with the latter offering at least some reassurance about their count, given the special circumstances surrounding dorm residents. Among the large group of "other noninstitutional" GQs,

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Table 9.4 Group Quarters by Occupancy Status and Type, 2020 Census

		Occupancy Status (%)		
GQ Type	Total GQs	Occupied	Vacant	Delete
Correctional (Adults)	17,000	79.4	18.8	1.8
Juvenile	9,000	73.3	24.4	2.2
Nursing Homes	29,000	89.7	9.0	1.6
Other Institutional	3,100	67.7	29.0	2.9
College/University Student Housing	39,000	84.6	13.8	2.1
Military	5,900	57.6	39.0	3.4
Other Noninstitutional	161,000	65.8	28.0	5.8
Total	264,000	72.3	23.3	4.4

NOTE: GQ, group quarters.

SOURCE: Tabulation from Census Unedited File (CUF), originally prepared for American Statistical Association Task Force on the Quality of the 2020 Census (WRF T2 ASADQS Table 11A All TEAs). See Disclosure Review Statement; CBDRB-FY23-0221.

including small group homes for adults and SBE locations, 28% were deemed vacant, with another 6% deleted during the enumeration. The level of deletions among "other noninstitutional" GQs was, by far, the highest for any major GQ type and, again, is likely evidence of the higher level of instability among addresses in this category, despite many being added just prior to or during the enumeration itself.

9.7.2 Imputation of Occupancy Status and Population Counts

As the 2020 Census was nearing completion, the Census Bureau recognized that many GQs had not provided the information necessary to determine occupancy status, much less population counts (Fontenot, 2022a). In response to this problem, the Census Bureau assembled a team to make final determinations regarding the occupancy status of GQs and, where appropriate, their population counts. The team conducted a telephone operation to contact GQs for which there was incomplete data. The results of this calling operation were then combined with information from the GQAC in January/February 2020, to determine occupancy status and population counts. (Recall that the GQAC operation collected the expected population and maximum population capacity expected on Census Day as well as notes from the interviewer regarding the GQ's operational status, including whether the GQ would be nonexistent or vacant on Census Day.) The GQAC information was especially important for those situations in which a population count of zero was indicated and a decision needed to be made about the accuracy of that count. In those situations in which the Census Bureau had a completed interview but no population

Copyright National Academy of Sciences. All rights reserved.

count, it was largely inferred that the GQ was vacant on Census Day. Moreover, due to the transient nature of GQs that provide services, such as shelters, soup kitchens, mobile food vans, and TNSOLs, the Census Bureau usually determined that a count of zero obtained during the enumeration was accurate and those GQs were deemed to be vacant.

For GQs that were determined to be occupied but with no population count, an imputation method provided a count. The method relied heavily on the expected and maximum counts from the GQAC operation in January/February 2020 and a series of indicators, called auxiliary data, which consisted of information from Census Bureau surveys (e.g., American Community Survey), various lists, and, in some cases, administrative data from other agencies (e.g., National Center for Education Statistics). If these sources were not available, the Census Bureau used data for the same GQ type within the same state as a basis for imputing counts. Thus, to the extent possible, the new GQ imputation process involved using other sources of data for the same GQ location and—barring that—information about different GQs of the same type.

An important feature of the imputation method was the use of ratios between the auxiliary information and the reported count for each GQ type within the state. The count from auxiliary data was then multiplied by the relevant ratio to get a final count for a GQ. For example, if the Census Bureau found that the count from a reporting hospital was, on average, 95% of the estimated count from their 2020 GQAC visits for all such hospitals across the state, a 2020 count was imputed for nonresponding hospitals that was 95% of the hospitals' estimated counts from 2020 GQAC (Fontenot, 2022a).

Of the 264,000 GQs, about 9% were deemed eligible for count imputation, with about one-quarter of those actually receiving imputed counts, which amounts to about 2% of all GQ facilities in the enumeration (Table 9.5). For the three largest GQ types by population, the percent of all GQs that were count imputed was very low for adult correctional facilities and nursing facilities (1.2%), but was well above the average for college/university student housing (3.1%). While "other noninstitutional" GQs had a count imputation rate slightly below the average for all GQs, these facilities accounted for more than one-half of all GQs with imputed counts—54%. The highest percentage of GQs with imputed counts was among two small groups—"other institutional" facilities (6.5%) and juvenile facilities (4.4%).

9.7.3 Methods of Data Collection by Group Quarters Type

The original operational plan for persons in GQs aimed to collect data in ways tailored to the characteristics of residents. In some cases, data were available electronically from records maintained by federal agencies (e.g., military, federal penitentiaries) or from administrators of facilities (e.g., state

Copyright National Academy of Sciences. All rights reserved.

Table 9.5 Count Imputation for Group Quarters, 2020 Census

GQ Type	Total GQs	Eligible for Count Imputation		Eligible and Count Imputed			Percent Distribution		
		N	Percent of All GQs	*N*	As Percent of Eligible GQs	As Percent of All GQs	All GQs	Eligible GQs	Eligible and Count Imputed
Total Eligible GQs	264,000	23,000	8.7	5,500	23.9	2.1	100.0	100.0	99.1
Correctional (Adults)	17,000	1,500	8.8	200	13.3	1.2	6.4	6.5	3.6
Juvenile	9,000	1,200	13.3	400	33.3	4.4	3.4	5.2	7.3
Nursing Homes	29,000	800	2.8	350	43.8	1.2	11.0	3.5	6.4
Other Institutional	3,100	500	16.1	200	40.0	6.5	1.2	2.2	3.6
College/University Student Housing	39,000	2,000	5.1	1,200	60.0	3.1	14.8	8.7	21.8
Military	5,900	500	8.5	100	20.0	1.7	2.2	2.2	1.8
Other Noninstitutional	161,000	16,500	10.2	3,000	18.2	1.9	61.0	71.7	54.5

NOTES: GQ, group quarters.

SOURCE: U.S. Census Bureau - RFI21.3 GQ_CI_NI_GQtype. See Disclosure Review Statement; CBDRB-FY23-0221.

Copyright National Academy of Sciences. All rights reserved.

prisons and other correctional facilities). For many GQ types, however, the options available for data collection were far more varied, for example for skilled nursing facilities, where records are usually kept for licensing and/or reporting purposes, but where access can be an issue, depending on local recordkeeping requirements and methods. In other cases, such as the collection of data for persons experiencing homelessness, collection usually involves fieldwork at places where services are obtained, such as shelters and mobile food van sites.

The Census Bureau temporarily suspended field operations on March 18, 2020, in response to the emerging COVID-19 pandemic. As the pandemic continued, it became clear that the Census Bureau needed to adjust its plans for enumerating GQs. This was especially the case for three populations: those living in college/university student housing, who left campus early and were no longer present to complete a form; and nursing homes and correctional facilities, where restrictions on access frustrated data-collection efforts. In these cases, the Census Bureau contacted facility administrators who originally opted for in-person enumeration, to encourage the submission of information electronically. In-person visits to GQs did not resume until July 1, 2020. Finally, the Census Bureau issued a letter to college and university presidents on June 15,[5] noting the importance of counting college students "where they usually live as of April 1, 2020, even if they left town early because of a school closure or shift to distance learning." However, the letter mentioned the GQ effort—collecting data on the on-campus student housing population—only in passing; instead, the letter's principal focus was the off-campus student population, living in housing units near campus that are not owned by the college or university. The letter advised the schools that the Census Bureau would soon request their assistance in obtaining "basic information about students who live off-campus, including local address information," by sending "records through a secure portal or by other arrangements."

Promoting electronic response on the college/university student housing population (both the GQ on-campus and regular housing unit off-campus segments) was a necessary measure given the COVID-19 pandemic conditions. However, it had the effect of exacerbating a known and long-standing mismatch between census law (in Title 13 of the U.S. Code) and the Family Educational Rights and Privacy Act (FERPA; codified at 20 U.S.C. § 1232g). In short, absent direct interview of students (and hence expressed consent), FERPA prohibits colleges and universities from releasing personally identifiable information from education records with very few exceptions, the decennial census not being defined as such an exception. However, FERPA does permit the release of "directory information" without express consent, provided that a student has not explicitly opted out of such disclosure. The FERPA language itself defines

[5]See https://www.census.gov/content/dam/Census/newsroom/press-kits/2020/addp-2l-students-2020-letter.pdf.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

"directory information" as including "name, address, telephone listing, [and] date and place of birth," as well as such data as major field of study and dates of attendance at the school; however, the definition ultimately depends on what information the college or university designates as "directory information" through public notice. In August 2019, the Census Bureau circulated a letter to student housing administrators reiterating its commitment to collect data in compliance with FERPA, foreshadowing its intent to request "permitted disclosure under FERPA" of directory information "for students who do not respond;" on January 29, 2020, the U.S. Department of Education circulated a detailed letter to all institutions of higher education reiterating its interpretation of FERPA as it applies to the 2020 Census, explicitly noting that it has, "since at least 1991, . . . consistently advised that the disclosure of a student's sex, ethnicity, or race as directory information would not comply with the regulatory definition."[6] In short, the Census Bureau found itself in a position of having to push for increased electronic submission of at least *some* information on college students, even though FERPA generally prohibits the sharing of key census data items through "directory information" submitted directly by the college or university.

In light of the COVID-19 pandemic, the need to temporarily abandon efforts to conduct the SBE—at places including temporary shelters, soup kitchens, mobile food van locations, and certain outdoor locations—also became apparent. Because it is intended to represent the service-receiving population of persons without a permanent residence as of the Census Day reference point, SBE was planned to occur similarly to 2010—with deployment of field staff to conduct the count on or very near Census Day (in this case, scheduled for March 30–April 1, 2020). After consulting with stakeholders and experts on how best to conduct the SBE on a different date (and in a different season of the year), SBE operations were moved to September 22–29, 2020. The Census Bureau closely coordinated visits to shelters and TNSOLs with administrators and local officials in advance, to facilitate the count of persons experiencing homelessness.

For most GQ types, quality checks were performed. Cases were assigned for reinterviews and a special form was used to follow up. Area Census Office staff used the reinterview form to telephone each GQ and confirm that the enumerator visited the correct facility at the correct address. The reinterview determined if the enumerator obtained a population count within an acceptable range (defined as plus or minus 20%) of the count provided by the GQ contact during reinterview.

[6]See https://studentprivacy.ed.gov/sites/default/files/resource_document/file/ED_SPPO%20Letter%20to%20IHEs%20re%202020%20Census_v_01292020.pdf; the August 2019 Census Bureau letter is copied as an attachment to the U.S. Department of Education letter.

Copyright National Academy of Sciences. All rights reserved.

As data collection proceeded, gaps and inconsistencies in the information appeared, in the form of partial responses and missing information. In addition, for college/university student housing, the Census Bureau observed that students were being reported in more than one location on campus, for example in a main location *and* a specific dormitory building. This observation prompted the Census Bureau to reinitiate contact with more than 10,000 GQs to collect additional information and to reconcile and/or confirm information that was already collected. This operation occurred from December 10–21, 2020, and involved a cadre of staff in field offices, the National Processing Center, and Census Bureau headquarters. When information was still missing, the Census Bureau used count imputation to fill in the gaps, which was the first time that count imputation was deployed for the GQ population.

As mentioned earlier, the Census Bureau collected data for persons in GQs using methods that differed based on the population to be enumerated and their residential living arrangements. The manner in which each enumeration mode was used in the 2020 Census also depended on the effects of the COVID-19 pandemic on residents. A broad outline of the methods used for selected GQ types is presented in Table 9.6, which provides a list of the methods, with notation indicating necessary changes implemented by the Census Bureau based on pandemic-related restrictions and logistical problems.

In-Person Enumerations occurred through face-to-face interviews using Individual Census Questionnaires administered by census fieldworkers. Originally, this operation was to take place in the field from April 1–June 12, 2020, in the United States and Puerto Rico. With the onset of the COVID-19 pandemic, however, the operation was postponed and took place from July 1–August 26, 2020. It is important to note that in-person enumeration was the sole source of data for persons enumerated in SBE locations other than shelters, including soup kitchens, mobile food vans, and in TNSOLs.

Drop-Off/Pick Up involved enumerators setting a time to drop off paper questionnaires to a GQ administrator or other contact person, as well as a time to pick up the completed questionnaires. Using a paper roster, the enumerator compared the number of questionnaires collected with the number of residents listed on the roster. If the numbers differed, the enumerator reached out to the GQ administrator or other contact person for a resolution.

Facility Self-Enumeration involved a Census Bureau-sworn GQ representative (e.g., facility staff member) conducting the enumeration of residents at the GQ. This method was used in correctional facilities, military installations, and medical facilities. In some cases, other GQ facilities requested to self-enumerate their residents, given the risks posed by the COVID-19 pandemic and potential disruption to their operations.

Paper Response Data Collection (PRDC), sometimes referred to as *Paper Listing*, involves enumerators picking up a template containing resident response information, specifically, first and last name, age, date of birth, ethnicity, race,

Copyright National Academy of Sciences. All rights reserved.

Table 9.6 Allowable Enumeration Methods by Group Quarters Type and Code, 2020 Census

	In-Person Interview	Drop Off/ Pick Up	Facility Self-Enumeration	Paper Response Data Collection	eResponse
Federal Correctional Facilities (101, 102)[a]	No	No	No	No	Yes
State, Local Prisons (103, 104, 105); Juvenile Facilities (202, 203); Other Institutional (401, 402, 403)	Yes	Yes	Yes	Yes	Yes
Military Jails (106)	No	No	Yes	No	Yes
Juvenile Group Homes (201); College/University Student Housing (501, 502); Adult Group Homes (801); Workers' Group Living Quarters (901)	Yes	Yes	Yes[b]	Yes	Yes
Skilled Nursing Facilities (301); Adult Residential Treatment (802)	Yes	Yes	Yes	Yes	Yes
Military Treatment Facilities (404); Military Quarters (601)	No	Yes[b]	Yes	Yes[b]	Yes[b]
Shelters (701)	Yes	Yes[b]	No	Yes	No
Soup Kitchens, Mobile Food Vans (702, 704)	Yes	No	No	No	No
Targeted Non-Sheltered Outdoor Locations (706)	Yes	No	No	No	No

[a] Drop Off/Pick Up, Facility Self-Enumeration, Paper Response Data Collection, and eResponse were all offered to code 101 and 102 correctional facilities operated by the Bureau of Indian Affairs.

[b] These modes were not initially offered to these GQ types, but were made available when operations were altered due to the COVID-19 pandemic.

SOURCE: Adapted from Group Quarters Operations Integrated Project Team, "2020 Census Group Quarters Operational Assessment Study Plan—Final." Version 1.22, June 8, 2022. See Disclosure Review Statement; use of this table was included in CBDRB-FY23-0221.

Copyright National Academy of Sciences. All rights reserved.

and alternate address information. After undergoing a series of checks at the Area Census Office level, the lists were sent to the National Processing Center, where clerks entered response information for the GQ into the Decennial Tracking and Control System (DTCS).

In addition, the Census Bureau offered an *alternate Paper Response Data Collection* method, in which GQ contact persons could mail back paper listings to Area Census Offices, which would then send the information to the National Processing Center for entry into the DTCS. This method was especially useful in situations in which COVID-19-related restrictions discouraged face-to-face contact. PRDC templates were used for these two operations from April 20–August 26, 2020.

The *eResponse* option provided facility administrators with a means of submitting electronic records to the Census Bureau via the Secure Web Incoming Module portal, which was developed to permit transfer of geographic data from local partners to the Census Bureau. The point of contact for the facility was given a unique user ID to electronically submit response data for each person living or staying at the facility on Census Day. The electronic portal provided viewable and printable information, including: a message to GQ administrators encouraging participation in the census; a confidentiality statement; the standard template (in Microsoft Excel format) and a sample entry; and a toll-free help desk number. Data were expected to be entered into the Excel template provided by the Census Bureau and uploaded to the portal; when submitted in the expected manner, GQ administrators were notified that their data had been received, and also notified of any missing records or errors in the data received. Submissions that did not follow the standard template format were sent to the National Processing Center and treated as paper data-collection responses. eResponse was the sole method used for the collection of data for federal correctional facilities, based on data from the Federal Bureau of Prisons. With the onset of the COVID-19 pandemic, the Census Bureau extended eResponse availability through August 28, 2020, and urged facilities to utilize eResponse whenever practical.

The large majority of all GQs in the 2020 Census provided data to the Census Bureau via eResponse and Paper Listings. Almost half of all GQs used eResponse and about one-third used Paper Listings (Table 9.7). Another 11% provided data via Facility Self-Enumeration. A submission by an administrator, based on facility records, is *not* considered to be an administrative records enumeration, but part of the general data collection; hence the very small number in the Administrative Records column.[7] The HU to GQ, UHE Processing column includes enumerations initially thought to be housing units,

[7]This category consists largely of administrative records enumerations of what were originally deemed housing units but were converted to GQs, and records of the deployed military. In addition, some records submitted through the Off-Campus Administrative Records Program for what were deemed GQs are included.

Copyright National Academy of Sciences. All rights reserved.

Table 9.7 Persons in Group Quarters by Group Quarters Type and Method of Data Collection, 2020 Census

	Method of Data Collection (%)							
GQ Type	Overall	eResponse	Paper Listing	Facility Self-Enumeration	Administrative Records	HU to GQ, UHE Processing	Other Operational Sources	Count Imputation
Total	8,238,969	49.0	33.5	10.7	0.0	1.0	3.7	2.0
Correctional (Adults)	1,960,127	76.8	17.7	2.2	0.0	0.1	2.2	0.9
Juvenile Facilities	87,252	46.0	32.3	11.9	0.0	1.0	3.2	5.6
Nursing Homes	1,622,249	37.4	42.0	14.1	0.0	0.4	4.8	1.2
Other Institutional	70,894	33.2	40.7	10.4	0.0	0.9	6.2	8.7
College/University Housing	2,790,556	58.1	32.7	1.6	0.1	0.3	4.4	2.8
Military	328,833	4.3	71.7	14.1	0.2	0.2	7.0	2.5
Other Noninstitutional	1,379,058	16.6	38.2	36.4	0.0	4.5	2.1	2.2

NOTES: GQ, group quarters; HU, housing unit; UHE, usual home elsewhere.

SOURCE: WRF T2 ASADQS Table 11A All TEAs. See Disclosure Review Statement; CBDRB-FY23-0221.

Copyright National Academy of Sciences. All rights reserved.

but that turned out to be GQs, as well as processing for persons with Usual Home Elsewhere (UHE) resulting in persons counted in a GQ instead of a housing unit. This accounted for about 1% of all persons in GQs. Other Operational Sources of GQ counts consisted of counts collected through the GQ Calling Operation, and the redistribution of facility-level counts to a single GQ. Concerning the latter, it is common for individual GQs (e.g., dormitories, hospital buildings, military barracks), to be part of larger GQ facilities. In some instances, the Census Bureau had to redistribute *total* facility-level counts that were inadvertently assigned to a single building, leaving other buildings without populations (e.g., total hospital facility count assigned to a single GQ building on a medical campus). Another similar problem occurred when some universities submitted counts for *all* students in college/university student housing for *each* dorm on campus, creating duplication that necessitated correction (Fontenot, 2022a; Thieme, 2021).

As emphasized in the earlier discussion, overall numbers for GQs can be very deceiving as a measure for any one type or category. For the three largest GQ categories, the methods used to collect data varied substantially. Correctional facilities made maximum use of eResponse (77%), likely the result of the requirement for federal facilities to participate via that route, unlike those at state and local levels, for which other response modes were permitted. Among responses from college/university student housing, reliance on eResponse also accounted for a majority of all responses (58%); but, given the wider variety of response options available, about one-third used Paper Listings as a response pathway. For the third largest GQ group—residents of nursing homes—the picture is even more varied. With all response mode options available, 42% responded with Paper Listings, 37% via eResponse, and an additional 14% used Facility Self-Enumeration.

For those in other noninstitutional GQs, response options also varied. Those in group homes for adults and in shelters for persons without a permanent residence had a fairly wide set of response options available; but, for those in soup kitchens, mobile food van locations, and TNSOLs, the only available mode was an in-person interview. Thus, Paper Listings and Facility Self-Enumeration accounted for the majority of enumerations. At the other end of the spectrum were military GQs, for which the large majority (72%) of responses were collected via Paper Listings, and Facility Self-Enumeration accounted for 14%. Originally, to enumerate military personnel in GQs, the Census Bureau planned to use administrative records through the Defense Manpower Data Center. After initial testing demonstrated that these files were faulty, the Census Bureau decided to permit other response pathways, the primary one being a paper-based reporting using Military Census Reports (Fontenot, 2021g:4).

Copyright National Academy of Sciences. All rights reserved.

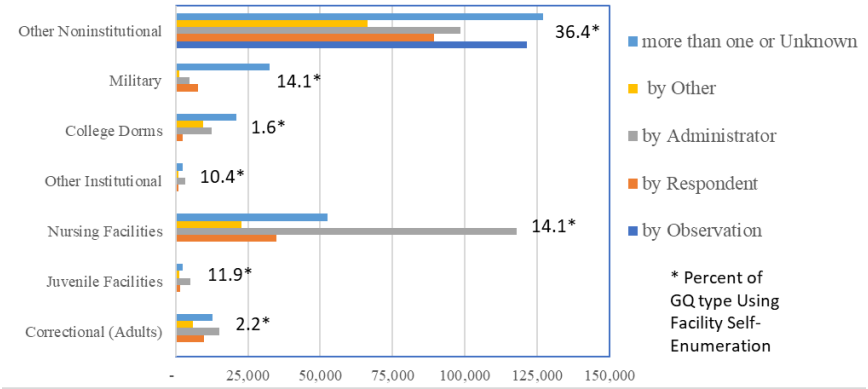


Figure 9.1 Person data collected using Facility Self-Enumeration by group quarters type and method of collection, 2020 Census.

SOURCE: WRF T2 ASADQS Table 11A All TEAs. See Disclosure Review Statement; CBDRB-FY23-0221.

9.7.4 Facility Self-Enumeration

As discussed earlier, Facility Self-Enumeration involved a Census Bureau-sworn GQ representative (e.g., facility staff member) conducting an enumeration of residents at the GQ. Depending on the living arrangements and circumstances of residents, differing personnel or methods could be used to collect the characteristics data within a facility: through observation, administrator queries, or interview with respondents themselves. Figure 9.1 shows how Facility Self-Enumeration was carried out within each of the GQ types. The large and varied mix of intake methods for those in "other noninstitutional" GQs is a testimony to the variety of subgroups in this category. This is quite significant because more than 36% of all data collection in "other noninstitutional" GQs made use of this method. The other observation that stands out is the substantial contribution of facility administrators in the data collection for nursing facilities. While comprising just 14% of all data collection for such places, the data obtained through administrators was likely pivotal, given the difficulties encountered in these facilities as a result of the COVID-19 pandemic.

9.7.5 Item Nonresponse by Group Quarters Type

Beyond the counts themselves, collecting data on the characteristics of GQ residents proved to be difficult for a number of GQ types in the 2020 Census

Copyright National Academy of Sciences. All rights reserved.

Table 9.8 2020 Census Item Nonresponse Rates for Group Quarters Person-Level Items, by Type

GQ Type	Sex (%)	Age/Date of Birth (%)	Hispanic Origin (%)	Race (%)
All GQs	22.0	17.8	43.9	30.2
Institutional GQs	4.7	4.8	27.1	8.7
Correctional Facilities for Adults	3.2	3.4	25.9	5.0
Juvenile Facilities	4.8	5.8	22.8	10.8
Nursing Facilities	6.4	6.2	28.7	12.9
Other Institutional Facilities	9.9	9.4	34.0	16.1
Noninstitutional GQs	36.7	28.9	58.0	48.4
College/University Student Housing	43.9	27.2	62.4	54.3
Military Quarters	25.6	40.5	76.6	62.8
Other Noninstitutional Facilities	24.9	29.4	45.3	33.5

NOTE: GQ, group quarters.

SOURCE: 2020 Census Unedited File, tabulation by DeJesus and Konya (2023:16).

Box 9.3 Imputation of Group Quarters Characteristics

The U.S. Census Bureau imputed values for missing characteristics first and foremost by using data from other information on a group quarters (GQ) roster for that person, in what is called a "joint edit." For example, if Hispanic origin is missing but race is reported, Hispanic origin is filled in following the same procedure as for household data, using what can be gleaned from the race response. Other data may be reported elsewhere, for example, for people enumerated in two places (e.g., college students mistakenly enumerated at their parents' home), and the Census Bureau can use characteristics data from the duplicate response if values are missing from GQ data. The Census Bureau can make use of previous census data, information for a respondent collected in the American Community Survey, or administrative data from the Social Security Numident file. Finally, if none of these options yield the necessary information, hot deck imputation is used to create a response, stratifying the "donor" GQ population by GQ type, age of respondents, and geographic area.

(Table 9.8). For all GQs, the level of missing data was substantial, across the board by characteristic: age (18%), sex (22%), race (30%), and Hispanic origin (44%). All of these rates represent substantial increases over rates seen in 2010 (Figure 9.2). For the missing characteristics the Census Bureau developed an imputation method to fill the void (see Box 9.3 for a description of imputation for GQ characteristics).

Like so many observations involving GQs, however, overall numbers can be deceiving and only begin to tell the story of missing data. In general, institutional GQs had much lower rates of missing information on each of the four key variables than did noninstitutional GQs. While it is impossible to reach any definitive conclusions with the available data,

Copyright National Academy of Sciences. All rights reserved.

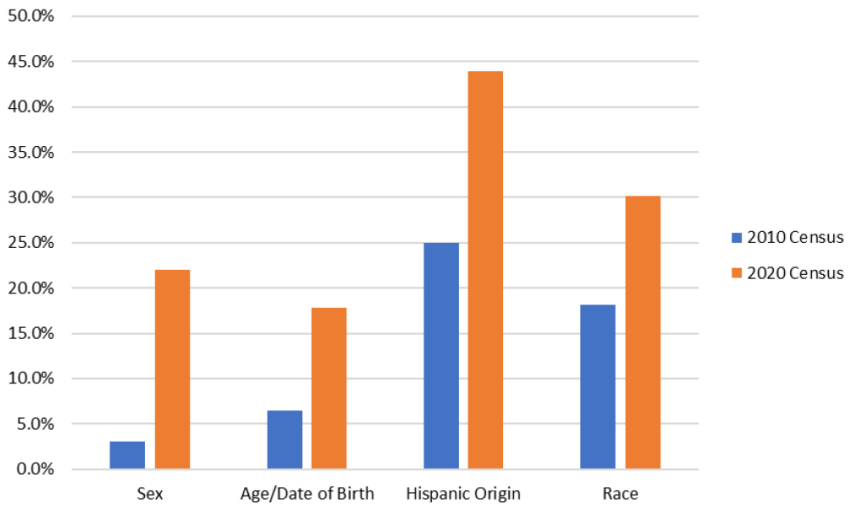


Figure 9.2 2010 and 2020 Census item nonresponse rate comparison of person-level items, group quarters enumerations.

SOURCE: 2020 Census Unedited File, tabulation by DeJesus and Konya (2023:49).

it is likely that the differences between institutional and noninstitutional GQs relate to the less stable nature of populations in these specific settings. Residents in noninstitutional GQs come and go more freely, which may make capturing their characteristics more difficult, as with college students living in college/university student housing, persons in small group homes for adults, and those in service-based locations.

Just 5% of residents in institutional GQs were missing data on sex and age, while the percentages were 37 and 29%, respectively, for noninstitutional GQs. For race, the picture was similar, albeit more pronounced—9% for persons in institutional settings, compared with 48% for those in noninstitutional settings. While higher nonresponse was also characteristic of reporting on Hispanic origin for those in noninstitutional GQs (54%) compared to institutional GQs (27%), even more important was the very high level for all GQ types. For specific types of GQs, two observations stand out: First is the high level of missing data for sex and age for the three noninstitutional GQ types—no less than one-quarter of residents missing this very basic information. Especially noteworthy are the college/university enumerations missing data on sex and the military enumerations missing information on age. Second are the ultra-high levels of missing data on race and Hispanic origin for these

Copyright National Academy of Sciences. All rights reserved.

same noninstitutional GQ types, with the military having the highest levels of missing data for Hispanic origin (77%) and race (63%). As a next step, it is instructive to ask whether patterns of missing data may bear some relationship to the mode of data collection. To this end, we now examine item nonresponse by GQ type by mode of data collection, with a focus on the two modes that accounted for the largest part of the data collection: eResponse and Paper Listings.

eResponse

The effort to automate data collection using eReponse is laudable and represents an important innovation, especially given the circumstances surrounding the COVID-19 pandemic. In general, missing data rates were low for all items obtained via eResponse for most GQ types, a very important and instructive finding given the high overall rates of missing data for GQs (Figure 9.3). For sex, age, and race, rates of missing data for most GQ types that reported via eResponse were well under 5%. Rates of missing data for Hispanic origin were higher for "other institutional" GQs, but only for adult correctional facilities was the rate higher than 10%, modest by overall standards. This is especially important for those large-category GQ types that made heavy use of eResponse, such as adult correctional facilities, of which 76% utilized this mode, and among nursing facilities, of which close to 40% used eResponse. Perhaps the automated nature of eResponse and programmed completeness checks paid off in the form of better reporting.

The one big exception to this encouraging pattern occurred for college/university student housing, for which eResponse was used heavily but the FERPA restrictions on sharing anything but "directory information" (see Section 9.7.3) resulted in very high rates of missing data for basic items: age (16%), sex (27%), race (32%) and Hispanic origin (35%).

Reporting of Hispanic Origin by State: A Closer Look at College/University Student Housing Using eResponse

To understand where the higher rates of missing data for college/university student housing GQs had the most impact, Table 9.9 shows the states with the highest levels of missing data for the question on Hispanic origin. Of those states with the highest percentages missing for Hispanic origin, all used eResponse well in excess of the 49% average for all states. Delaware, the District of Columbia, and Arizona stand above the pack, with rates of missing data of at least 70% and utilization of eResponse between 63–77%—high by any standard and, as a result, very impactful on overall levels of missing data on this attribute. An examination of the level of missing data for college/university student housing on the Hispanic-origin question for all states reveals a moderate positive

Copyright National Academy of Sciences. All rights reserved.

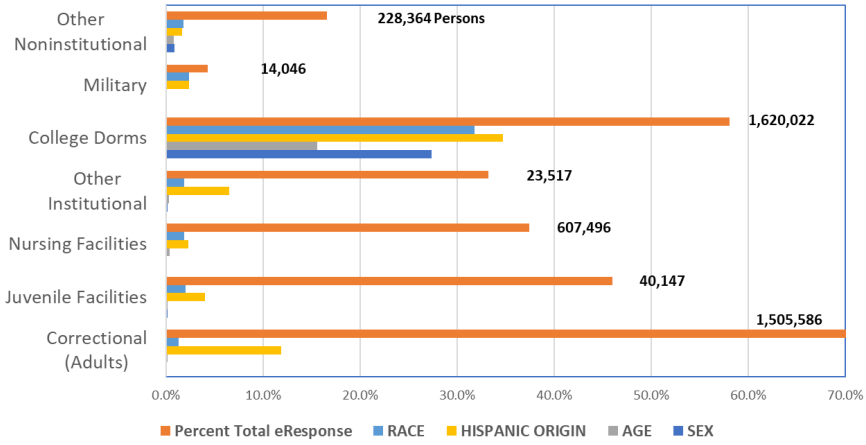


Figure 9.3 Item nonresponse rates by group quarters type, data collection via eResponse.

SOURCE: WRF10 GQ INR Rates by Response Type. See Disclosure Review Statement; CBDRB-FY23-0221.

relationship ($r = 0.48$) with the total utilization of eResponse (Figure 9.4). This begs the question of whether this relationship is somehow related to the Hispanic-origin composition of a state, given the concerns about answering the census in Hispanic communities throughout the nation. Figure 9.5 shows that there is little relationship between percent Hispanic origin in a state and the level of missing data for Hispanic origin ($r = 0.09$). It is important to remember, however that this analysis is for the population in college/university student housing whose information was collected via eResponse, a subset of the population that may be far different from the overall Hispanic-origin composition at the state level.

The unique results for college/university student housing, relative to other GQ types, certainly imply that the deficiency in the reporting of missing data is related to respondent resistance and/or confusion over the proper interpretation of FERPA. However, more research is needed to examine why the eResponse system failed to detect and/or correct this deficiency.[8] Further, work needs to be done at executive and legislative levels to clarify reporting methods for institutions of higher education, given the FERPA requirements,

[8]Attention should also be paid to possible effects of high rates of missing data items for off-campus college students, to the extent to which such data were gathered through eResponse-type directory information from institutions of higher education.

Copyright National Academy of Sciences. All rights reserved.

Table 9.9 Hispanic-Origin Item Nonreponse Rates, eResponse for College/University Student Housing, by State

State	College/ University Housing Responses (all modes)	Total eResponse	Percent Total eResponse	Hispanic Item Nonresponse, in eResponse
Delaware	11,147	6,970	62.5	78.6
District of Columbia	23,839	18,340	76.9	72.4
Arizona	38,948	26,890	69.0	69.8
New Hampshire	22,792	16,808	73.7	56.8
Alabama	45,486	31,428	69.1	56.2
Oregon	25,398	17,801	70.1	52.6
Georgia	82,345	52,536	63.8	52.2
Rhode Island	25,945	16,694	64.3	49.1
Maryland	46,124	30,715	66.6	49.0
Wisconsin	60,151	49,468	82.2	48.9
Ohio	111,714	78,013	69.8	48.5

NOTE: Table is limited to states with approximately 49% or more nonresponse to the Hispanic-origin item.

SOURCE: WRF10 GQ INR Rates by Response Type. See Disclosure Review Statement; CBDRB-FY23-0221.

and to clarify operational improvements in time for the 2030 Census. Given the level of automation, it should be possible to conduct checks on the eResponse system to further drill down and, with some qualitative interviews with college administrators, try to determine why levels of missing data were so high for college/university student housing in some states and not in others.

Paper Listings

As discussed earlier, the Paper Listings mode involves the use of a template that is filled out by GQ administrators and reviewed at a number of points by census workers in Area Census Offices and in the National Processing Center, as the data from a facility or location are entered into the system. With the exception of college/university student housing, data collection via Paper Listings was subject to higher levels of missing data on characteristics across GQ types, relative to data collected via eResponse (see Figure 9.6).

More than any other GQ type, those in military GQs had high levels of missing characteristics data, which is especially important because of the high reliance of this GQ type on Paper Listings (70%). The high item nonresponse for Hispanic origin was 69% and helps explain the high overall item nonresponse for Hispanic origin of 77%, discussed earlier. The elevated

Copyright National Academy of Sciences. All rights reserved.

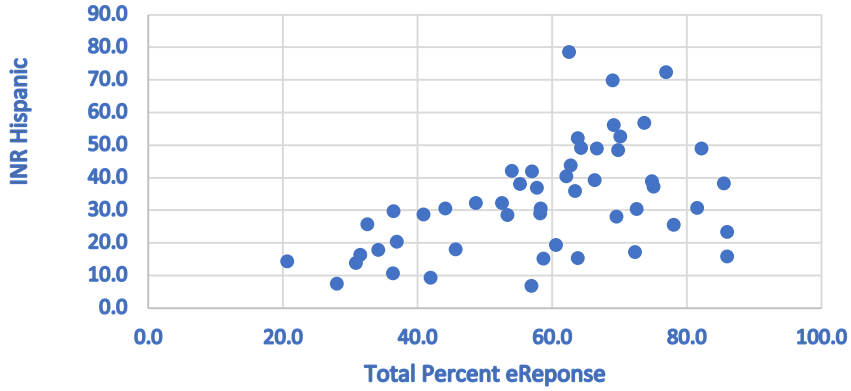


Figure 9.4 Hispanic-origin item nonresponse rates by percent responding by eResponse, college/university student housing, 50 states and District of Columbia.

NOTE: INR, item nonresponse.

SOURCE: Source: WRF10 GQ INR Rates by Response Type. See Disclosure Review Statement; CBDRB-FY23-0221.

level of missing data on the other three items for military GQs is indicative of a data-collection anomaly that needs to be addressed for 2030—especially given the breakdown of quality-assurance checks that were supposed to occur during the Paper Listing template-intake process.

The fact that college/university student housing has elevated levels of missing information on characteristics via the Paper Listings mode lends credence to the idea that this problem relates more to the rules regarding administrative data access than to the mode of collection. Again, a remedy that addresses the rules regarding FERPA needs to be developed for the 2030 Census.

Finally, two other observations are instructive. The first is the inability of Paper Listings to capture data on Hispanic origin, relative to other characteristics, across the board by GQ type. The pervasiveness of this finding necessitates further research, to understand the underpinnings of the apparent reticence on this attribute. Did something in the mechanics of reporting or an attribute of the environment produce this shortfall? A second observation concerns the completeness of data for "other noninstitutional" GQs, which in this case includes a large number of adult group homes, residential treatment facilities for adults, and shelters for those experiencing homelessness—three ever-changing populations that can be particularly challenging to enumerate.

Copyright National Academy of Sciences. All rights reserved.

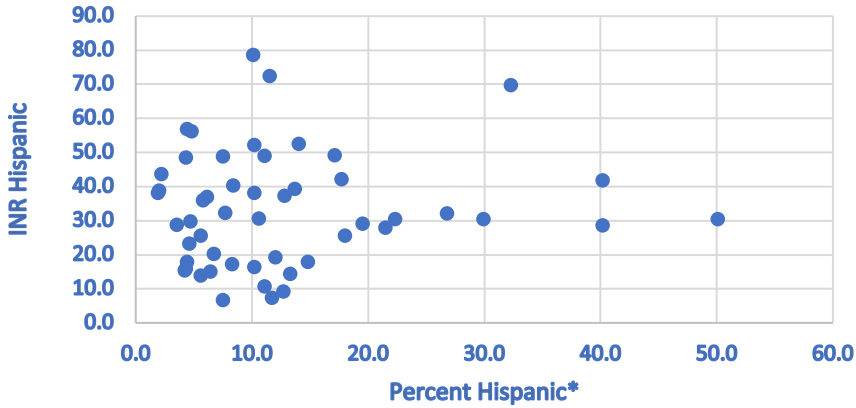


Figure 9.5 Hispanic-origin item nonresponse rates by percent Hispanic, college/university student housing, 50 states and District of Columbia.

NOTES: INR, item nonresponse. Percent Hispanic is derived from 2021-vintage population estimates by the U.S. Census Bureau.

SOURCE: Source: WRF10 GQ INR Rates by Response Type. See Disclosure Review Statement; CBDRB-FY23-0221.

Yet in this case, missing data levels were modest compared with military GQs and college/university student housing, and in the same range for several institutional GQ types.

9.8 CONCLUSIONS AND RECOMMENDATIONS

The diversity of living situations falling under the broad heading of the nonhousehold GQ population makes it a perennial challenge in the decennial census. GQ enumeration is always difficult, and the COVID-19 pandemic-induced operational adjustments were particularly hard to grapple with in this regard. As with other parts of the 2020 Census, Census Bureau staff rendered laudable and extraordinary efforts to adapt to the conditions and make the GQ count as complete as possible. Yet it is only fair to reiterate that not all of the 2020 Census challenges in GQ were unique to the circumstances of 2020.

> *Conclusion 9.1: The enumeration of nonhousehold group quarters (GQ) locations in the 2020 Census experienced major shortcomings, some of which can be directly attributed to COVID-19 pandemic adjustments to the data-collection scheduling (e.g., the scuttling of*

Copyright National Academy of Sciences. All rights reserved.

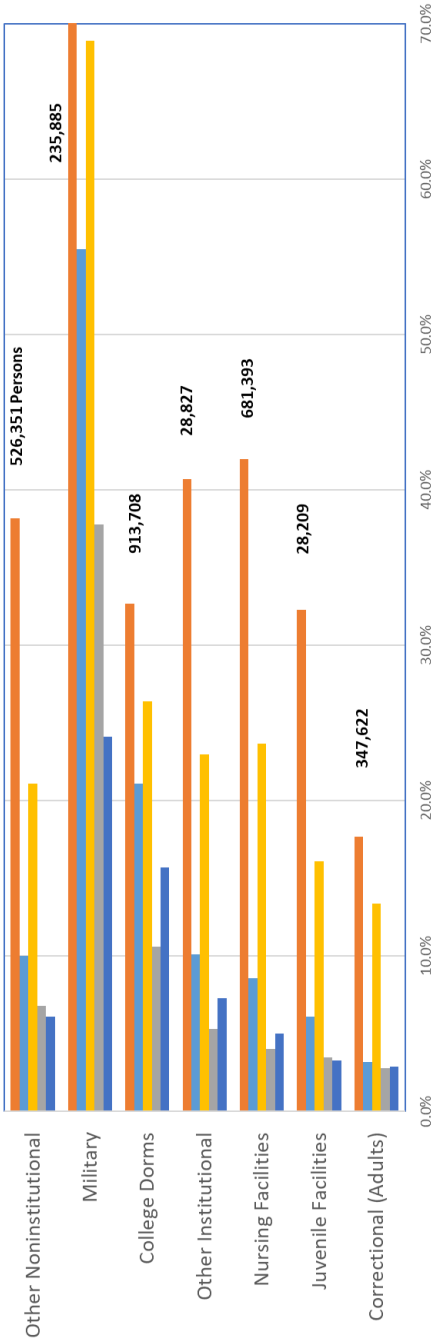


Figure 9.6 Item nonresponse rates by group quarters type, data collection via Paper Listing.

SOURCE: WRF10 GQ INR Rates by Response Type. See Disclosure Review Statement; CBDRB-FY23-0221.

Copyright National Academy of Sciences. All rights reserved.

> *prearranged plans for in-person enumeration worked out in the Group Quarters Advance Visit operation). Yet several important operational and conceptual shortcomings cannot be attributed solely to the pandemic, such as the mismatch between census content and the U.S. Department of Education's interpretation of permissible data provision under federal privacy law, the shortfalls in the enumeration of people in military GQs, as well as the fluctuation in status (and resulting problems in the address list) in more fluid GQ types like group homes.*

Indeed, in pondering root causes and opportunities for improvement, we find ourselves in alignment with a predecessor National Academies of Sciences, Engineering, and Medicine panel. Seventeen years ago, the Panel on Residence Rules in the Decennial Census looked back on the conduct of GQ enumeration in the 2000 Census and found the situation "unacceptably bad. Failure to reconcile the group quarters roster with the MAF contributed to a host of census errors. Group quarters frames were constructed without sufficient standardization and awareness of diversity in housing unit and group quarters stock," and some GQ data (particularly those generated from long-form questionnaire information from a sample) were "marred by extremely high levels of nonresponse" (National Research Council, 2006:Finding 7.1). That panel argued that the prevailing notion that this relatively small but very important segment of the population can be enumerated using exactly the same methods and approaches is "fundamentally flawed;" the living situations falling under the broad GQ heading are too diverse for a single methodology to be viable.

Preparations for the 2030 Census could benefit from an approach that acknowledges the reality that the GQ enumeration is not a monolithic entity, but a number of very different, distinct censuses, each with unique requirements. Consequently, no single precensus program can be expected to capture *all* addresses needed to conduct a complete enumeration of the GQ population, just as no single approach and contact strategy is completely satisfactory.

In some instances, GQs are stable, longstanding facilities, as with large correctional facilities, nursing homes, college/university student housing, and military bases, for which addresses are likely stable over time, and thus changes can be easily documented. Still, these need to be treated differently because of the unique enumeration issues, involving mismatches with other parts of privacy law (e.g., FERPA for college/university student housing and the Health Insurance Portability and Accountability Act of 1996 for nursing homes) and quality of facility records data (e.g., military bases).

In other cases, addresses for GQ respondents may be fleeting at best, and can only be "discovered" just before or during the enumeration. Such is the

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

case with more transient GQ locations, such as small group homes, residential treatment facilities, and places where those without a usual residence receive services. For this class of GQs, agility—in the form of flexible listing and enumeration procedures—needs to be the order of the day. Therefore, while it is important to curate a list of GQs throughout the decade, with some equivalent of an ongoing GQAC program, it is unrealistic to expect that such an activity alone will provide adequate coverage of the entire GQ universe for 2030.

For the 2010 Census, the Census Bureau hosted a group of local officials who had Special Sworn Status and access to local data, which permitted the Census Bureau to fix problems *before* release of the apportionment counts by December 31. For 2020, that review process was almost nonexistent, given the necessary data-collection delays but unchanging data-delivery deadline. The prerelease quality check process should be resurrected for the 2030 Census, given the malleable nature of certain GQ types and the exigencies surrounding their accurate enumeration.

To its credit, the Census Bureau promoted the use of eResponse when circumstances permitted, a beneficial option given the COVID-19 pandemic. Almost half of all persons in GQs were enumerated using eResponse. Moreover, with the exception of college/university student housing, eResponse yielded relatively low item nonresponse rates, which was of great benefit to those GQ types that made substantial use of that pathway, such as adult correctional facilities and nursing homes. The beneficial attributes of eResponse make extension of this data-collection mode a logical choice for 2030 Census data collection, wherever practical. Thus, in the panel's view, the Census Bureau should continue to explore ways to increase utilization of this pathway.

The high level of missing data for characteristics for persons in college/university student housing is likely tied to the reliance on administrative data for students, given their physical absence from student housing as a result of the COVID-19 pandemic. Differences in the level of directory information that was available, combined with the rules governing its release (i.e., FERPA), likely hampered the use of such data for enumeration purposes; and this was the case irrespective of data-collection mode. This begs for a combination of administrative, regulatory, and/or legislative solutions that clarify the use of these data for enumeration purposes.

Paper Listing data collection was subject to higher levels of item nonresponse, especially for persons in military GQs and those in college/university student housing. While FERPA is likely a reason for the latter, the remarkably high levels of missing data for the military, most of which were collected via Paper Listings, necessitates additional research to discover a way to effectively capture existing data that elude the current system. In general, more work is needed to determine why the checks built into the overall Paper Listing capture process failed to detect what are substantial deficiencies in reporting, especially

Copyright National Academy of Sciences. All rights reserved.

on Hispanic origin and to a lesser degree race, and in some instances, even sex and age.

Recommendation 9.1: The U.S. Census Bureau's steps to make the Master Address File an ongoing, continuous inventory of all living quarters—including group quarters (GQs)—should continue, with an eye toward developing updated information for GQs on par with the regular updating of conventional housing units. This should include a separate inventory of GQ types that have longstanding address provenance and those that need to be tracked more closely because they are continually changing.

Recommendation 9.2: Group quarters (GQs) require more than updated address information; they require continually updated contact strategies as well. The U.S. Census Bureau should develop "selected advance contact" protocols with GQ facilities by type (and their parent companies, as appropriate), akin to follow-ups performed with businesses, and state, local, and tribal partners should be engaged in this update activity.

Recommendation 9.3: In light of the 2020 experience, the U.S. Census Bureau should convene group quarters (GQ) stakeholders in a discussion of more effective and useful options for electronic provision of data for GQ residents, and on the issues involved in obtaining access to administrative records for GQ facilities. The completeness of such records should also be assessed by GQ type.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

– 10 –

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. In this chapter, we provide a brief history of race and ethnicity measurement in the decennial census, assess the quality of the race and ethnicity data for 2020, and make recommendations for 2030. To provide a comprehensive picture of race and ethnicity measurement in the 2020 Census, the chapter excerpts findings from other chapters that address specific quality measures. Quality attributes covered in this chapter include:

- Effects of changes in the race question format, data capture, and coding conventions for 2020 on race reporting—compared with 2010, these changes appeared to contribute to a substantial increase in the multiracial population and a decrease in the White Alone group;
- Coverage errors in the census for major race and ethnic population groups (see Chapter 4 for more detail);
- Rates of imputations for missing and inconsistent responses to the race and ethnicity questions; and
- Effects of the use of differential privacy-based algorithms to protect the confidentiality of 2020 Census responses on the timeliness and accuracy of race and ethnicity data (see Chapter 11 for more detail).

It is important to note that "race" and "ethnicity" are social-political constructs and often associated with national origins. The concepts are also

Copyright National Academy of Sciences. All rights reserved.

fluid, depending in part on the format of the survey or census questions[1] and who is the respondent (i.e., self, other household member, proxy).

10.1 OVERVIEW OF RACE AND ETHNICITY MEASUREMENT

Measuring race has been a core function of the constitutionally mandated U.S. decennial census since 1790 (that census identified White and all other). According to the U.S. Census Bureau's web site:[2]

> Information on race is required for many Federal programs and is critical in making policy decisions, particularly for civil rights. States use these data to meet legislative redistricting principles. Race data also are used to promote equal employment opportunities and to assess racial disparities in health and environmental risks.

The 1970 Census was the first census to include a specific question on ethnicity, or what was often termed Spanish origin, in its 5-percent long-form sample.[3] The 1980 Census and subsequent censuses asked ethnicity of everyone.

10.1.1 U.S. Office of Management and Budget Standards and Census Race and Ethnicity Questions

The 1977 and 1997 OMB Standards

In 1977, the Bureau of the Budget (now the U.S. Office of Management and Budget, OMB) issued Statistical Policy Directive No. 15, which specified standard race and ethnicity categories for federal use government wide. The legal authority for the directive stems from the 1942 Federal Reports Act, which mandated a "forms clearance process" for all forms used by the federal government to standardize classifications, avoid duplication, and lessen respondent burden. The Paperwork Reduction Act of 1980 and its reauthorizations in 1986 and 1995 updated and expanded the authority of OMB—specifically through the office of a chief statistician—to coordinate the federal statistical system and develop and oversee the implementation of

[1] A prime example of format likely influencing response on race and ethnicity is the 2000 Census question on Hispanic origin which, unlike the question used in 1990, 2010, and 2020, did not list specific examples in the "Other Spanish/Hispanic/Latino" category to prompt people in providing write-in answers. (The 2020 version provided six examples, such as Salvadoran and Dominican.) Responding to user concerns that the lack of examples led respondents to provide a generic response, such as "Spanish," and thereby undercount specific groups, Census Bureau researchers conducted analysis that supported this hypothesis (Cresce and Ramirez, 2003).

[2] See "About the Topic of Race," https://www.census.gov/topics/population/race/about.html.

[3] Prior censuses used different ways of estimating ethnicity, such as estimates of people of Puerto Rican origin in states and metropolitan areas with sufficiently large numbers of such people, and people with Spanish surnames in the Southwest.

Copyright National Academy of Sciences. All rights reserved.

Table 10.1 U.S. Office of Management and Budget Standard Race and Ethnicity Categories in Statistical Policy Directive No. 15

1977—Separate Questions Preferred	**1997**—Separate Questions Preferred
Race Categories:	*Race Categories:*
White	White
Black	Black
American Indian or Alaska Native (AIAN)	AIAN
Asian or Pacific Islander	Asian
	Native Hawaiian or Other Pacific Islander (NHOPI)
Ethnicity Categories:	*Ethnicity Categories:*
Hispanic	Hispanic/Latino
Not Hispanic	Not Hispanic/Latino

SOURCE: 1997, https://www.bls.gov/bls/statistical-policy-directive-15.pdf. 1977, https://obamawhitehouse.archives.gov/omb/fedreg_race-ethnicity, Appendix.

standards and guidance (see National Academies of Sciences, Engineering, and Medicine, 2021:App. A).[4]

OMB issued an update of Directive No. 15 in 1997 which, for the first time, stipulated that respondents should have the option to check more than one race category. The directive permits agencies to include additional race categories as long as they can be aggregated to the principal categories. The Census Bureau has traditionally collected expanded detail for race and ethnicity beyond that described in the OMB standard. "Other" is not a listed category in Directive No. 15, but Congress has required the Census Bureau to include "Some Other" in the decennial census and the American Community Survey.[5] Table 10.1 provides the OMB categories as of 1997 along with the 1977 categories for comparison.

Census Race and Ethnicity Questions, 1970–2020

In addition to the major change in the census race question in 2000 to a "check more than one" format, each census has seen minor changes in format, wording, and other features. Table 10.2 provides the categories for race (10.2(a)) and ethnicity (10.2(b)), respectively, in the 1970–2020 Censuses.

[4]The chief statistician heads the Statistical and Science Policy Office in the Office of Information and Regulatory Affairs, established by the 1980 Paperwork Reduction Act, in OMB.

[5]In 2009, Congress mandated in appropriations language that "none of the funds provided in this or any other Act for any fiscal year may be used for the collection of census data on race identification that does not include 'some other race' as a category" (123 Stat. 3115), in response to a proposal to delete the category from the 2010 Census questionnaire. The statutory notes

Copyright National Academy of Sciences. All rights reserved.

Table 10.2 Response Categories to Race and Ethnicity Questions, 1970–2020 Censuses

(a) Race Question

1970	1980	1990	2000 (mark one or more races)	2010 (mark one or more races)	2020 (mark one or more races *and* print origins)
White	White	White	White	White	White (print origins)
Negro or Black	Black or Negro	Black or Negro	Black, African American or Negro	Black, African American or Negro	Black or African American (print origins)
Indian (American) (print tribe)	Indian (American) (print tribe) Eskimo Aleut	Indian (American) (print enrolled or principal tribe) Eskimo Aleut	American Indian or Alaska Native (print enrolled or principal tribe)	American Indian or Alaska Native (print enrolled or principal tribe)	American Indian or Alaska Native (print enrolled or principal tribes)
Chinese Japanese Filipino Korean	Chinese Japanese Filipino Korean Vietnamese Asian Indian	Chinese Japanese Filipino Korean Vietnamese Asian Indian	Chinese Japanese Filipino Korean Vietnamese Asian Indian Other Asian (print race)	Chinese Japanese Filipino Korean Vietnamese Asian Indian Other Asian (print race)	Chinese Japanese Filipino Korean Vietnamese Asian Indian Other Asian (print origins)
Hawaiian	Hawaiian Guamanian Samoan	Hawaiian Guamanian Samoan Other Asian or Pacific Islander (print race)	Native Hawaiian Guamanian or Chamorro Samoan Other Pacific Islander (print race)	Native Hawaiian Guamanian or Chamorro Samoan Other Pacific Islander (print race)	Native Hawaiian Chamorro Samoan Other Pacific Islander (print origins)
Other (print race)	Other (specify)	Other race (print race)	Some other race (print race)	Some other race (print race)	Some other race (print race/origin)

SOURCE: Inspection of questionnaires, various online sites.

Copyright National Academy of Sciences. All rights reserved.

Table 10.2 (cont.)

(b) Ethnicity Question

1970 (5% sample long form only)	1980 (Hispanic question after race)	1990 (Hispanic question after race)	2000 (Hispanic question first)	2010 (Hispanic question first)	2020 (Hispanic question first)
No, none of these (last category shown)	No, not Spanish/Hispanic	No, not Spanish/Hispanic	No, not Spanish/ Hispanic/Latino	Not Hispanic, Latino or Spanish origin	Not Hispanic, Latino or Spanish origin
Mexican	Mexican, Mexican Am., or Chicano	Mexican, Mexican Am., or Chicano	Mexican, Mexican Am., or Chicano	Mexican, Mexican Am., or Chicano	Mexican, Mexican Am., or Chicano
Puerto Rican	Puerto Rican	Puerto Rican	Puerto Rican	Puerto Rican	Puerto Rican
Cuban	Cuban	Cuban	Cuban	Cuban	Cuban
Other Spanish	Other Spanish/Hispanic	Other Spanish/Hispanic (print one group)	Other Spanish/ Hispanic/Latino (print group)	Other Hispanic, Latino or Spanish origin (print origin)	Other Hispanic, Latino or Spanish origin (print origin)
Central or South American	—	—	—	—	—

SOURCE: Inspection of questionnaires, various online sites.

Copyright National Academy of Sciences. All rights reserved.

Research conducted as early as 1997 and throughout the next two decades, culminating in the 2015 National Content Test, demonstrated that a single combined race/ethnicity, check-more-than-one question would be preferable to the two-question format in several respects. A single question would be more intelligible to respondents, reduce the extent of missing data, and practically eliminate the Some Other Race (SOR) category, which about 40% of Hispanic people have typically checked as their race (Mathews et al., 2017). The 2015 National Content Test also considered the effect of adding a separate Middle Eastern or North African (MENA) race category, finding that people reporting relevant national or ethnic backgrounds chose the MENA category in a reinterview when it was available, and that reporting of MENA responses in the Some Other Race category decreased from 12% without a MENA category to 3% with a MENA category (Mathews et al., 2017).

In September 2016, OMB issued a *Federal Register* notice (81 FR 67398) requesting comments on four aspects of Statistical Policy Directive No. 15: (1) moving to a single combined check-more-than-one race and ethnicity question with Hispanic as a category; (2) the addition of a separate MENA race category, which had been considered for the 1997 revision but not approved pending further research; (3) clarifying that agencies are not limited to the minimum set of race and ethnicity categories; and (4) clarifying race and ethnicity terminology to keep up with changing societal usage. The notice identified a range of uses of federal data on race and ethnicity and the need to consider effects on data quality and utility of proposed changes to the categories and question format (see Box 10.1).

The Census Bureau expected that OMB would approve the move to a single question and the addition of a separate MENA race category. OMB never gave its approval, however, and the Census Bureau announced in December 2017 that it would continue with the two-question format without a MENA category.[6]

Format and Processing Changes to the 2020 Race and Ethnicity Questions

The Census Bureau decided to make significant format and processing changes to the 2020 race and ethnicity questions. These changes included:

- Adding space to write in origins under the White and Black checkboxes in addition to the other checkbox categories in the race question;[7]
- Expanding data capture of write-in responses from only 30 characters in 2010 to 200 characters in 2020;

accompanying 13 U.S.C. § 5, on the nature of Census Bureau questionnaires, indicate that similar language was inserted into appropriations acts earlier in 2009, as well as in 2007, 2005, and 2004.

[6]MENA responses are currently treated as White in accordance with the current OMB race category definitions.

[7]See Marks and Rios-Vargas (2021) and U.S. Census Bureau (2021d:Appendix B, Definitions of Subject Characteristics).

Copyright National Academy of Sciences. All rights reserved.

Box 10.1 Request for Comment on Proposed Race and Ethnicity Measurement Standards

Federal Uses of Race and Ethnicity Data: When providing comment regarding proposed areas for possible revision, it may be helpful to keep in mind how the standard is used. The standard not only guides information collected and presented from the decennial census and numerous other statistical collections, but also is used by Federal agencies for civil rights enforcement and for program administrative reporting. These include, among others:

- Enforcing the requirements of the Voting Rights Act;
- reviewing State congressional redistricting plans;
- collecting and presenting population and population characteristics data, labor force data, education data, and vital and health statistics;
- establishing and evaluating Federal affirmative action plans and evaluating affirmative action and discrimination in employment in the private sector;
- monitoring the access of minorities to home mortgage loans under the Home Mortgage Disclosure Act;
- enforcing the Equal Credit Opportunity Act;
- monitoring and enforcing desegregation plans in the public schools;
- assisting minority businesses under the minority business development programs; and
- monitoring and enforcing the Fair Housing Act.

To most effectively promote information quality, the intended uses of data on race and ethnicity should be considered when changes to the standards are contemplated. Additionally, the possible effects of any proposed changes on the quality and utility of the resulting data must be considered.

SOURCE: Excerpted from "Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity," 81 FR 67398, September 30, 2016.

- Expanding the coding of write-ins from only two detailed categories in 2010 to up to six detailed categories in 2020;
- Expanding the list of codes for 2020 based on input from experts on specific race and ethnicity categories;
- Combining the race and ethnicity code lists; and
- Coding from left to right, instead of prioritizing Hispanic responses for the ethnicity question and race responses for the race question as was done in 2010.

Looking specifically at the Hispanic-origin question, for 2020, the Census Bureau decided to code people as Hispanic if they responded "No, not Hispanic" but provided an Hispanic category (e.g., Mexican) in one of the race question write-in spaces. The Census Bureau would also code people as not Hispanic if they checked "Other Hispanic" in the Hispanic-origin question but

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

provided only non-Hispanic write-ins (this edit was also done in 2000 and 2010). Both edits overrode the actual response to the Hispanic-origin question.[8]

Looking at the race question, the Census Bureau assigned people to the specific race(s) corresponding to each check box they marked, even if write-ins did not support the categorization. The Census Bureau might assign other specific races depending on write-ins, including assigning a race from a write-in when the person did not check the corresponding box. For example, people who checked White and wrote in an American Indian tribe but did not check the American Indian or Alaska Native (AIAN) box were coded White and AIAN, while people who checked White and wrote in an Hispanic origin (e.g., Mexican) but did not check the SOR box were coded White and SOR. Write-ins were coded the same regardless of where they were written—for example, a write-in of Jamaican was coded as Black no matter where it was written (e.g., under White or SOR). Write-ins under the SOR checkbox were used where possible to assign one or more specific races and delete the SOR categorization.

A New OMB Standard?

In January 2023, OMB issued a *Federal Register* notice calling for public comments on proposed updates to Statistical Policy Directive No. 15, with the intent of finalizing a new standard by summer 2024 (88 FR 5375, January 27, 2023).[9] The core proposals, suggested by an interagency working group, are to implement a single combined question for race and ethnicity, adding "Hispanic or Latino" and MENA as new main categories; the proposal also invites comment on the feasibility of requiring detailed race and ethnicity categories by default (i.e., check-box examples and write-in responses) rather than the minimum, main category headings. As of August 2023, the notice had drawn more than 20,000 comments.

10.1.2 Race and Ethnicity Distributions from 1970–2020—A More Diverse Nation

Figure 10.1 graphs the distributions of race and ethnicity from 1970–2020. The bottom six groups (from White Alone to Two or More Races) add up to 100% from responses to the race question. The top group is the Hispanic population, from responses to the separate ethnicity question—Hispanics may be of any race. Table 10.3 tabulates the percentage distribution of race and ethnicity from 1970–2020 for the non-Hispanic population (top

[8]This paragraph and the next rely on notes from Jeffrey Passel, Pew Hispanic Research Center, based on his conversations with Census Bureau staff.

[9]See also the Census Bureau's notice of the public comment period at https://www2.census.gov/about/ombraceethnicityitwg/omb-proposals-english.pdf. The *Federal Register* notice requested comments by April 27, 2023.

Copyright National Academy of Sciences. All rights reserved.

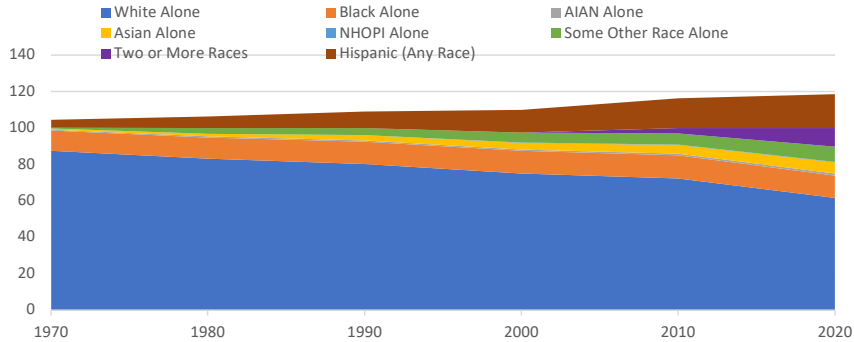


Figure 10.1 Percentage distribution of the population by race and ethnicity, 1970–2020 Censuses.

NOTES: AIAN, American Indian or Alaska Native; NHOPI, Native Hawaiian or Other Pacific Islander.

SOURCES: 1970: Gibson and Jung (2002:Table 1) and U.S. Census Bureau (1973:Table 2). 1980: Tabulation at https://censusscope.org/us/chart_race.html. 1990: U.S. Census Bureau (2001:Table 4). 2000: Grieco and Cassidy (2001:Table 10). 2010, 2020: U.S. Census Bureau (2021g:Table 4).

half of table) and Hispanic population (bottom half of table). Focusing on changes from 2010–2020, the White Alone category decreased somewhat among non-Hispanic people and markedly among Hispanic people (denoted in red). Conversely, the Two or More Races category increased somewhat among non-Hispanic people and substantially among Hispanic people (denoted in green).

By age, younger people were more diverse in race and ethnicity than people ages 18 and over in both 2010 and 2020, while racial and ethnic diversity increased between 2010 and 2020 for both age groups (see Figure 10.2). Specifically, Figure 10.2(a) shows the increase in the Hispanic population over time and for younger people compared with older people, while Figure 10.2(b) similarly shows the decrease in the White Alone population and the increase in the Two or More Races population. Additional age detail from the 2020 Census to examine racial and ethnic diversity among generations was not available until the release of the 2020 Demographic and Housing Characteristics file in May 2023 (see Chapter 11 and Section 10.5).

Copyright National Academy of Sciences. All rights reserved.

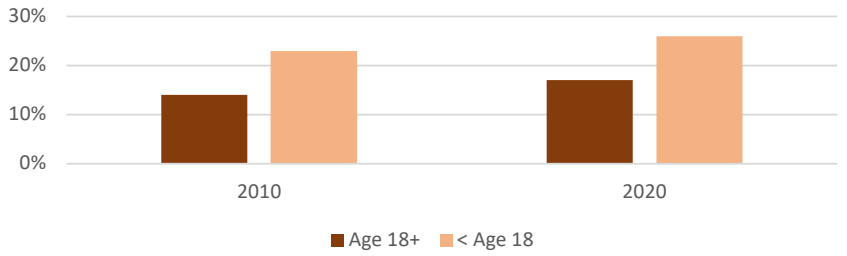
Table 10.3 Percentage Non-Hispanic/Hispanic of Total Population and Percentage Distribution of the Non-Hispanic/Hispanic Population by Race, 1970–2020 Censuses

Census/Category	1970	1980	1990	2000	2010	2020
Non-Hispanic Population (1,000s)	194,229	211,933	226,356	246,116	258,268	269,369
Percent Non-Hispanic of Total (%)	95.5	93.6	91.0	87.5	83.7	81.3
Percent of Non-Hispanic:	100.0	100.0	100.0	100.0	100.0	100.0
White Alone	—	85.1	83.1	79.1	76.2	71.2
Black Alone	—	12.5	12.9	13.8	14.6	14.8
AIAN Alone	—	0.7	0.8	0.8	0.9	0.8
Asian/NHOPI Alone	—	1.6	3.1	4.2	5.8	7.5
Some Other Race Alone	—	0.3	0.1	0.2	0.2	0.6
Two or More Races	N.A.	N.A.	N.A.	1.9	2.3	5.0
Hispanic Population (1,000s)	9,073	14,609	22,354	35,306	50,478	62,080
Percent Hispanic of Total (%)	4.5	6.4	9.0	12.5	16.3	18.7
Percent of Hispanic:	100.0	100.0	100.0	100.0	100.0	100.0
White Alone	93.3	—	51.7	47.9	53.0	20.3
Black Alone	5.0	—	3.4	2.0	2.5	1.9
AIAN Alone	0.3	—	0.7	1.2	1.4	2.4
Asian/NHOPI Alone	0.4	—	1.4	0.4	0.5	0.5
Some Other Race Alone	1.0	—	42.7	42.2	36.7	42.2
Two or More Races	N.A.	N.A.	N.A.	6.3	6.0	32.7

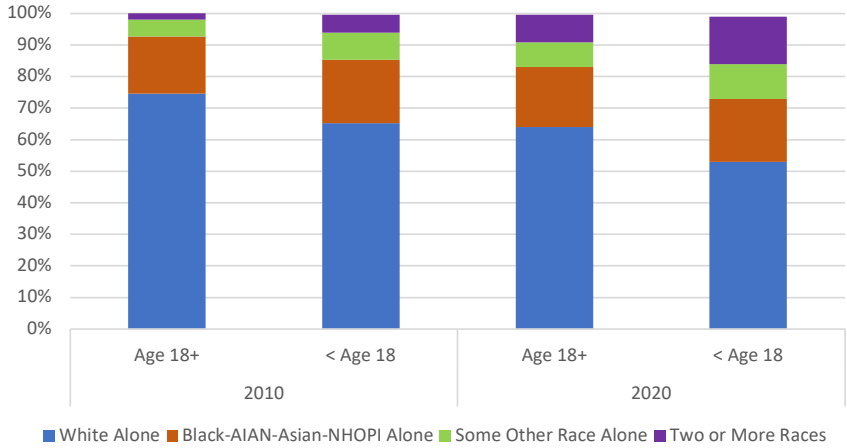
NOTES: AIAN, American Indian or Alaska Native; NHOPI, Native Hawaiian or Other Pacific Islander; N.A., not available; —, could not locate or not published. Major decreases from 2010 to 2020 are rendered in red, major increases from 2010 to 2020 are rendered in green.

SOURCES: 1970: Gibson and Jung (2002:Table 1) and U.S. Census Bureau (1973:Table 2). 1980: Tabulation at https://censusscope.org/us/chart_race.html. 1990: U.S. Census Bureau (2001:Table 4). 2000: Grieco and Cassidy (2001:Table 10). 2010, 2020: U.S. Census Bureau (2021g:Table 4).

Copyright National Academy of Sciences. All rights reserved.



(a) Percentage Hispanic



(b) Percentage distribution by race

Figure 10.2 Percentage Hispanic and percentage distribution by race for people over and under age 18, 2010 and 2020 Censuses.

NOTES: AIAN, American Indian or Alaska Native; NHOPI, Native Hawaiian or Other Pacific Islander.

SOURCES: 2020 and 2010 Redistricting Files, Tables P3, P4, accessed from data.census.gov.

Copyright National Academy of Sciences. All rights reserved.

10.2 NEW WRITE-INS, DATA CAPTURE, AND CODING FOR 2020

10.2.1 Effects on Race and Ethnicity Estimates from the Census and the American Community Survey

The provision for new write-ins and expanded data capture and coding undoubtedly contributed to the dramatic changes in the distribution of race in the 2020 Census compared with 2010—see Table 10.4, which also compares the 2020 American Community Survey (ACS) with the 2019 ACS.[10] There was a substantial drop in the White Alone population (shown in red) and a substantial increase in the Two or More Races population (shown in green) in both the 2020 Census and 2020 ACS, mostly but not entirely concentrated among people who answered Hispanic to the ethnicity question (see Table 10.3). The coding change that appears to account for this dramatic swing among Hispanic people was a combination of allowing write-ins under the White checkbox and coding any Hispanic entry in the White write-in space (or any other race category write-in space) as SOR, thereby recategorizing Hispanic people who only checked White (and who would have been White Alone in 2010 and previous censuses) as Two or More Races (White and SOR).

In addition to the substantial increase in Two or More Races among Hispanic people in the 2020 Census (20.3 million up from 3.0 million in 2010), there was an increase in Two or More Races among non-Hispanic people (13.5 million up from 6.0 million in 2010). The largest groups contributing to this increase of 7.5 million were:

- White and AIAN non-Hispanic people—increase of 2.3 million;
- White and SOR non-Hispanic people—increase of 2.1 million;
- White and Black non-Hispanic people—increase of 1.2 million; and
- White and Asian non-Hispanic people—increase of 1 million.

The increase in non-Hispanic White and AIAN people could be due to more people checking both White and AIAN. It could also be due to people who checked only White (and no other race including AIAN) and provided an AIAN origin in the White write-in space. Whether these people intended to provide two races or to provide a more detailed origin has not been analyzed. The increase in non-Hispanic White and SOR people does not have a ready explanation, although there are non-Hispanic SOR codes in the coding scheme, including Brazilian and Caribbean.[11]

[10]The 2020 ACS used the 2020 Census race and ethnicity format and coding, while the 2019 ACS used the 2010 Census format and coding.

[11]Brazilians, as noted, are supposed to be coded as non-Hispanic under the OMB 1997 standard, regardless of their responses to the Hispanic-origin question. Through an error in the 2020 ACS, Brazilians who said they were Hispanic were not recoded as non-Hispanic. This error also occurred for people who checked Hispanic and wrote in Belize or the Philippines. Consequently, 2020 ACS data show an out-of-proportion increase in Hispanics among these groups—most noticeably, fully

Copyright National Academy of Sciences. All rights reserved.

Table 10.4 Percentage Distribution of the Population by Race and Ethnicity, 2010–2020 Censuses and 2019–2020 1-Year American Community Survey Estimates

Census/ACS Year/ Percent in Category	2010 Census	2020 Census	2019 ACS	2020 ACS
Total Population (1,000s)	308,746	331,449	328,240	329,484
Percent of Total (%)				
White Alone	72.4	61.6	72.0	62.7
Black Alone	12.6	12.4	12.8	12.1
AIAN Alone	0.9	1.1	0.9	1.0
Asian Alone	4.8	6.0	5.7	5.7
NHOPI Alone	0.2	0.2	0.2	0.2
Some Other Race Alone	6.2	8.4	5.0	6.8
Two or More Races	2.9	10.2	3.4	11.5
Percent of Total (%)				
Hispanic	16.3	18.7	18.4	18.6
Non-Hispanic	83.7	81.3	81.6	81.4

NOTES: ACS, American Community Survey; AIAN, American Indian or Alaska Native; NHOPI, Native Hawaiian or Other Pacific Islander. Major decreases from 2010/2019 to 2020 are rendered in red, major increases from 2010/2019 to 2020 are rendered in green.

SOURCE: 2010, 2020 Censuses: U.S. Census Bureau (2021g:Table 1 (numbers), Table 2 (percentages)). 2019 ACS: Table B03002 (Hispanic or Latino Origin by Race), accessed via data.census.gov. 2020 ACS: 2020 ACS 1-Year Experimental Data Tables XK200201 (Race) and XK2003 (Hispanic origin) at https://www.census.gov/programs-surveys/acs/data/experimental-data/1-year.html.

10.2.2 Possible Causes

The Census Bureau has not produced tabulations for the same people (e.g., a sample of the 2020 records) using the 2010 and 2020 race question formats and data-capture and coding schemes, nor has it looked at whether the race identification of people included in both censuses remained the same or changed. Consequently, the degree to which the new schemes contributed to the increase in the Two or More Races category shown in Table 10.4, and the degree to which demographic change (e.g., more multiracial children) or changes in self-identification drove the increase is not knowable. The dramatic changes in the race distributions between the 2019 and 2020 ACS (similar in magnitude to those between the 2010 and 2020 Censuses) suggest that the new format (write-in spaces for the White and Black checkboxes) and the expanded data capture and coding played a major role.

70% of Brazilians were coded as Hispanic in the 2020 ACS compared with about 3–4% in the 2019 and 2021 ACS. This footnote was revised after release of the prepublication version of the report to correct the use of "White" versus "non-Hispanic" and more precisely reflect the findings in Passel and Krogstad (2023).

Copyright National Academy of Sciences. All rights reserved.

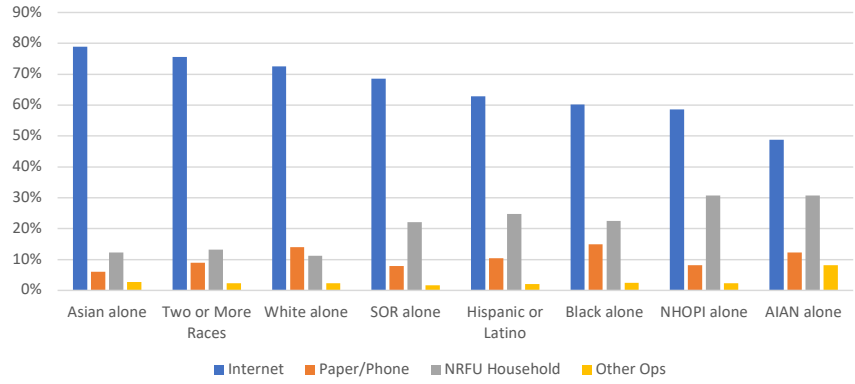


Figure 10.3 Percentage distribution of response modes for race and ethnic groups, 2020 Census.

NOTE: SOR, Some Other Race; NHOPI, Native Hawaiian or Other Pacific Islander; AIAN, American Indian or Alaska Native; NRFU, Nonresponse Followup. Includes people who reported their race and ethnicity themselves (or their households reported for them), excluding responses from proxies, imputations, and administrative records. Other Ops primarily includes smaller operations outside of Self-Response and NRFU responses. For the 2020 Census, this category included Update Enumerate, Remote Alaska, Enumeration of Transitory Locations, Coverage Improvement, and Self-Response Quality Assurance operations.

SOURCE: Census Bureau tabulations from 2020 Census Edited File. See Disclosure Review Statement; CBDRB-FY23-0179.

The Census Bureau provided the panel with limited tabulations that suggest areas for further research. Figure 10.3 shows the response modes in 2020 for people who provided their own race and ethnic identification (i.e., their responses were not obtained from a proxy interview in Nonresponse Followup (NRFU), imputed due to a missing response, or obtained from administrative records), by race and ethnic group.[12] The groups are ordered by the percentage who responded via the internet, from a high of 79% (non-Hispanic Asian Alone people) to a low of 49% (non-Hispanic AIAN Alone people). Among Hispanic people, 63% responded via the internet.

An examination of response modes by race and ethnicity illustrates that the propensity to write in responses varied by mode (see Figure 10.4).[13] Specifically, for every internet respondent who checked one or more major race category

[12]For many "self" reporters of race and ethnicity, the actual respondent was a member of the household who filled out the census questionnaire for all household members.

[13]Overall, there was an explosion of write-ins for the race question in 2020 compared with 2010—from almost 38 million in 2010 to almost 336 million in 2010 (write-ins under the Hispanic-

Copyright National Academy of Sciences. All rights reserved.

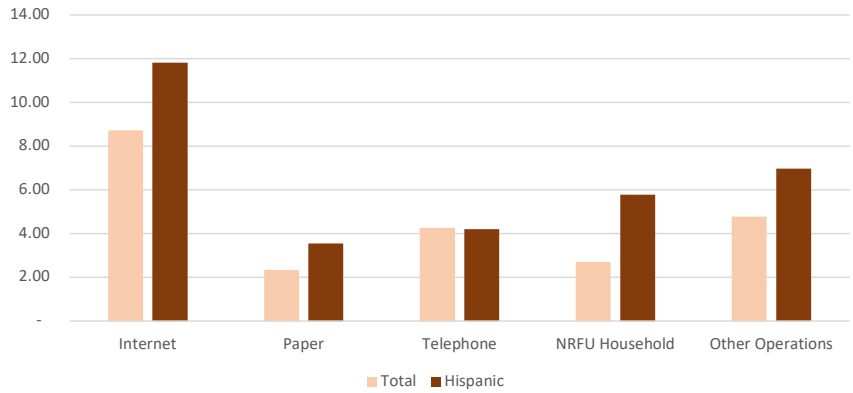


Figure 10.4 Ratio of write-in race responses (for people who provided at least one write-in) to checkbox-only responses by response mode, total and Hispanic population, 2020 Census.

NOTES: NRFU, Nonresponse Followup. Includes people who reported their race and ethnicity themselves (or their household reported for them), excluding responses from proxies, imputations, and administrative records. Checkbox-Only indicates that one or more major race categories were checked and no write-ins were provided. See also notes to Figure 10.3.

SOURCE: Census Bureau tabulations from 2020 Census Edited File. See Disclosure Review Statement; CBDRB-FY23-0179.

boxes but did not provide any write-ins, there were 8.7 write-ins by internet respondents who provided at least one write-in. For every Hispanic internet respondent who checked one or more major category boxes with no write-ins, there were 11.8 write-ins by internet respondents who provided at least one write-in. These ratios were lower for other response modes, although Hispanic people in every instance provided a larger number of write-ins than the total population.[14] Had Hispanic people responded via the internet at the same rate as the total population, then it is likely that the White Alone population would have declined even more and the Two or More Races population increased even more than what occurred in 2020.

origin question declined from 17 million in 2010 to 15 million in 2020). See Jones et al. (2021:Slide 9).

[14]For some people in large households, some paper questionnaires did not provide space to record race or ethnicity. Some write-in responses were removed through the editing process.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

10.3 COVERAGE ERROR

Every census misses people and at the same time duplicates people and makes other erroneous enumerations, the balance of which determines whether the census experienced a net undercount or overcount. More concerning, net undercounts and overcounts differ among population groups, which is why the paramount question about the race and ethnicity results in any census is the extent of differential net undercount. Results from the Census Bureau's two main methods to estimate coverage errors—a Post-Enumeration Survey (PES) and Demographic Analysis (DA)—were released in March–May 2022. The DA coverage error results by race (for Black people and all others) could not be released, however, because of delays due to implementation of the new 2020 Disclosure Avoidance System (DAS)—see Section 10.5.2.

Chapter 4 presents and discusses PES estimates of coverage errors in 2010 and 2020 by race and ethnicity cross-classified with other characteristics, including age, sex, and housing tenure (owner, renter). Summarizing the findings:

- Net *overcounts* were substantially higher in 2020 compared with 2010 for White people and Asian people, while net *undercounts* were substantially higher for people in the categories of Black, AIAN on reservations, Native Hawaiian or Other Pacific Islander (NHOPI), SOR, and Hispanic.
- The *differential coverage rate* (the difference in net coverage rates between two groups), which is critical for equitable allocation of legislative seats and federal and state funds, widened between non-Hispanic White Alone people and other groups from 2010–2020. Differences (from non-Hispanic White Alone people) increased from 2.9–4.9 percentage points for Black Alone or in Combination people, from 5.7–7.3 percentage points for AIAN Alone or in Combination people on Reservations, and from 2.4–6.6 percentage points for Hispanic people.
- Non-Hispanic White Alone people were almost always overcounted in both censuses, although the percentages are not large except for men and women ages 50 and over in rental housing, who were overcounted in 2020 by double the rate of 2010.
- Asian people were predominantly overcounted in both censuses but especially so if they lived in rental housing in 2020.
- Black people were generally undercounted in both censuses and particularly so if they lived in rental housing; net undercount rates were somewhat higher for this group in 2020 compared with 2010.
- Hispanic people under age 50 were generally undercounted in both censuses, with much higher rates in 2020 compared with 2010 and even higher rates if they lived in rental housing.

Copyright National Academy of Sciences. All rights reserved.

Table 10.5 Percentage of People with at Least One Item Reported Who Had Ethnicity Imputed and Percentage Who Had Race Imputed, by Response Mode, 2010–2020 Censuses

Census/Response Mode	2010 Ethnicity	2020 Ethnicity	2010 Race	2020 Race
Total Population (%)	4.5	8.7	4.1	9.2
Self-Response	4.8	2.4	4.2	2.8
Enumerator Return	3.3	26.8	3.7	29.0
People in Group Quarters	25.7	46.2	19.0	32.9

NOTES: Total Population excludes whole-person imputations for which only the existence of a person is known and not any of their characteristics; Self-Response includes internet, paper, and telephone; Enumerator Returns include NRFU and other operations such as Update Enumerate, and (for 2020) exclude administrative records enumerations.

SOURCE: DeJesus and Konya (2023:Table N).

Looking at components of error (duplications, other types of erroneous enumerations such as counting a baby born after Census Day, whole-person imputations, and omissions), the 2020 Census clearly experienced significant problems with the enumeration of Black people, AIAN people on reservations, and Hispanic people. There were also problems with the enumeration of NHOPI people. A striking finding is that the already-high rates of omissions for Hispanic people and those checking SOR Alone or in Combination (mostly Hispanic people) in 2010 (8.6 and 7.7%, respectively) were higher still in 2020 (9.9 and 10.5%, respectively).

10.4 IMPUTATIONS FOR MISSING AND INCONSISTENT DATA

Not every household provides answers for all items in the census questionnaire. Some responses also turn out to be inconsistent with other responses. The Census Bureau imputes values for missing and inconsistent responses based on other information for the same person or household (which could include a previous census or ACS response), termed "assignment," or by "allocating" a response from a person in a nearby household (or group quarters, GQ).[15] Table 10.5 provides 2010 and 2020 imputation rates for ethnicity and race by type of enumeration—self, enumerator, and GQ.[16]

[15]"Allocation" is typically referred to as "hot deck imputation" in the survey literature (see Citro, 2011).

[16]Whole-person imputations, which require imputing all characteristics and are not treated as item imputations, increased from 2% of the population in 2000 and 2010 to 3.4% of the population in 2020, primarily due to an increase in the percentage of households that only provided the population count (Khubba et al., 2022:Table 3).

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

The 2020 ethnicity and race item imputation rates for self-responses were lower in 2020 than in 2010, but they were higher overall and markedly so for enumerator returns and for people in GQs. Census tracts with low self-response, which are characterized by low income and low educational attainment among other characteristics, likely had high missing data rates for race and ethnicity and hence high imputation rates.[17] Overall, non-Hispanic Asian, White, and Two or More Races people had high self-response rates while other groups had lower rates and therefore likely higher imputation rates (see Figure 10.3). Looking at enumerator returns, proxy NRFU enumerations had higher imputation rates than other types of enumerator returns in both 2010 and 2020, and markedly higher rates in 2020 compared with 2010.[18]

To fill in missing race and ethnicity, the Census Bureau used one of the following assignment methods when possible: accept a Hispanic origin provided in the race question (for people missing ethnicity) or a race category provided in the ethnicity question (for people missing race); accept a person's 2010 Census or ACS response if such were available; or accept the race and Hispanic origin of another household member. If assignment was not feasible, as a last resort, the Census Bureau used allocation or hot deck imputation (essentially, accept the race and Hispanic origin of a neighbor). Assignments and allocations were used about equally for ethnicity and race imputations. For enumerator returns, which had high rates of imputation in 2020, 14% and 13% of ethnicity responses were assignments and allocations, respectively, as were 15% and 14% of race responses (DeJesus and Konya, 2023:Table E).

10.5 2020 DISCLOSURE AVOIDANCE SYSTEM (DAS)

The new DAS for 2020, using differential privacy-based algorithms (see Chapter 11), has had and continues to have significant adverse effects on the timeliness of release of 2020 data products. It also has forced the Census Bureau's DA program to delay release of net coverage results by race, and forced the population estimates program to continue to use race and ethnicity distributions derived from the 2010 Census, in what is termed a "blended base." Finally, the DAS introduced noise into small population groups and small governmental jurisdictions, which calls into question the usability of the data for important purposes.

[17] O'Hare and Lee (2021:Figure 2), using census tract characteristics from the 2015–2019 ACS.

[18] Based on comparisons of missing data rates (which are similar to but usually 1–3 percentage points lower than imputation rates). Proxy enumerations had missing data rates of 17% (ethnicity) and 15% (race) in 2010 compared with 38% (ethnicity) and 41% (race) in 2020 (U.S. Census Bureau, 2021c).

Copyright National Academy of Sciences. All rights reserved.

Box 10.2 Content of the 2020 Redistricting Data File

- Table P1: Race for total population—six major race categories alone, plus two races (15 combinations, such as White-Black, Black-American Indian and Alaska Native, etc.), three races (20 combinations), four races (15 combinations), five races (6 combinations), and six races, for a total of 63 race categories (71 cells including subtotals)
- Table P2: Hispanic origin and race for non-Hispanic people—same detail as total population (73 cells)
- Table P3: Race for total population ages 18 and older—same detail as total population (71 cells)
- Table P4: Hispanic origin and race for non-Hispanic people ages 18 and older—same detail as total population (73 cells)
- Table 5: Group quarters population by major type of group quarters (10 cells)
- Table 6: Occupancy status for housing units (3 cells)

SOURCE: Technical documentation accompanying 2020 Census State Redistricting Data (Public Law 94-171) Summary File.

10.5.1 Delays

A major consequence of the new DAS has been a series of unprecedented delays in the release of 2020 Census data products. Subject-matter and geographic detail have also been cut back in the 2020 data products compared with earlier censuses.

Regarding timing, between August 2021 and May 2023, the only dataset released to date (in addition to the 50 state population counts for reapportionment) was the 2020 P.L. 94-171 Redistricting File, released in August 2021. Box 10.2 lists its content, which essentially comprise race and ethnicity information for the total and voting-age population for all geographic units down to the block level.

In May 2023, the Census Bureau released the Demographic and Housing Characteristics (DHC) File, the successor to Summary File 1 (SF1) from previous censuses and a workhorse for data users. The DHC File contains cross-tabulations of race/ethnicity with sex, detailed age, housing tenure, and household relationship, with more detail available for higher levels of geography. The DHC File excludes some of the tables that were in SF1 or provides less geographic disaggregation for them—either due to confidentiality concerns or because the TopDown Algorithm used for the Redistricting File and the DHC File cannot handle complex household/family-person "joins," such as the number of households by average size, tenure, and race/ethnicity of the household head (see Chapter 11).

For the 2010 Census, SF1 was delivered on a flow basis, state by state, from June–August 2011 (a national file was released in October 2011), and earlier censuses met the same schedule. The 2020 DHC File was not released until May

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

25, 2023. This is unprecedentedly late, even allowing an additional 5–6 months because of the delay in census operations due to the COVID-19 pandemic. The 2020 DAS requires that a differential privacy-based algorithm be applied data product by data product, in contrast to previous censuses for which disclosure avoidance was implemented at the source.

Following release of the DHC File will be three files that use different types of differential privacy-based algorithms:[19]

- Detailed DHC-A (DDHC-A): Population counts and sex and age statistics for 300 detailed race and ethnic groups and 1,187 American Indian/Alaska Native/Native Hawaiian tribes and villages; released September 21, 2023.
- Detailed DHC-B (DDHC-B): Household type and tenure information for detailed race and ethnic groups and AIAN tribes and villages. As of this writing, DDHC-B is planned to be released in September 2024.
- Supplemental DHC (S-DHC): Household/family-person joins, including average household size by age and tenure, average family size, household/family type for children under 18 years and total population in households by tenure. Some tables will be iterated by major race and ethnicity categories (White Alone, Black Alone, etc.) As of this writing, S-DHC is planned to be released in September 2024.

DDHC-A and DDHC-B were collectively intended to be the equivalent of Summary File 2 (SF2) in previous censuses. The 2010 SF2 was released on a flow basis, state by state, from December 2011 through April 2012, with a national file released in May 2012. DDHC-A was not released until September 2023 and DDHC-B is not scheduled for release before September 2024. Moreover, the subject matter detail in both the A and B Files has been reduced compared with the 2010 SF2, and S-DHC will contain a greatly reduced number of person-household join variables compared with SF1 in previous censuses (see Chapter 11).

10.5.2 Effects of Delays on Race and Ethnicity for Demographic Analysis and Population Estimates

Although it would have been normal practice to release DA results by race (Black and all other races) in the same timeframe as the PES results, the April 2022 DA release only covered age and sex.[20] To permit race comparisons, 2020 Census race data need to be modified—specifically, people in the SOR Alone category need to have a specific race imputed for them because the DA estimates

[19]See the "About 2020 Census Data Products" listing at https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/release/about-2020-data-products.html.

[20]See U.S. Census Bureau (2022d).

Copyright National Academy of Sciences. All rights reserved.

derive from vital records and other sources that historically have not always had a SOR category. The Census Bureau therefore distributes SOR Alone responses to specific race groups (people classified as SOR and a specific race are assigned the specific race category and removed from the Two or More Races category) in a Modified Race File. The method used in previous censuses essentially searched for a donor—someone in the household or neighboring area who reported a specific race(s) and the same Hispanic origin (yes/no) as the person who checked SOR—and assigned the race(s) of the donor to that person.[21] Given the need to develop an appropriate differential privacy-based algorithm, the Census Bureau experienced delays in creating a 2020 Modified Race File on the same schedule as in 2010. The Census Bureau hoped to release the 2020 file in summer 2023 but has not done so as of this writing.

The lack of a Modified Race File for 2020 has also affected the annual population estimates, mandated in Title 13. These estimates provide vital inputs to federal formula allocations for states and localities, provide controls for the ACS and other household surveys, serve as denominators for vital rates and per-capita time series and as indicators of recent demographic changes, among many other uses.[22] The estimates typically derive from the most recent census, with race modified to allocate SOR Alone to specified races, updated with administrative records and survey data (birth and death records, Internal Revenue Service and Social Security data, Medicare records, ACS data, and other sources). Among the regularly produced series are estimates of single years of age by sex, race, and Hispanic origin for the nation, states, and counties.

The delays in the 2020 Census, concerns about data quality, and the delays in developing an appropriate differential privacy-based algorithm for the Modified Race File led to an unprecedented decision at the Census Bureau to continue the estimates going forward from the 2010 Census base, with some modifications (U.S. Census Bureau, 2021f). What the Census Bureau calls the "blended base" uses:

- 2020 Census population counts for the nation, state, and counties;
- Age and sex estimates from the 2020 DA (which does not exhibit the extreme age heaping evident in 2020 Census data and corrects for the large census undercount of young children—see Chapter 4) as a national control; and
- Race/ethnicity by age and sex for the nation, states, and counties from the *2010-based* population estimates updated through 2020 and forward.

Table 10.6 shows the 2020 population distribution from: (1) the 2020 population estimates based on updating the 2010 Census and released in 2020; (2) the 2020 population estimates using the blended base and released in 2021;

[21]For the 2020 method, see U.S. Census Bureau, Population Division (2012).
[22]See https://www.census.gov/programs-surveys/popest.html.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Table 10.6 Percentage Distribution of the Population by Race/Ethnicity, 2020 Population Estimates, 2020 Census

Source/ Race-Ethnicity Group	(1) 2020 Population Estimates (2010 Base)	(2) 2020 Population Estimates (July) (Blended Base)	(3) 2020 Census, as Reported (with Some Other Race)	(4) 2020 Census, with Some Other Race Reallocated
Total (%)	100.0	100.0	100.0	100.0
White Alone (%)	76.0	75.9	61.6	74.4
Black Alone	13.5	13.5	12.4	13.3
AIAN Alone	1.3	1.3	1.1	1.7
Asian Alone	6.3	6.1	6.0	6.3
NHOPI Alone	<0.1	0.3	0.2	0.3
SOR Alone	—	—	8.4	—
Two or More Races	2.9	2.9	10.2	4.0
Hispanic	18.6	18.7	18.7	18.7

NOTES: AIAN, American Indian or Alaska Native; NHOPI, Native Hawaiian or Other Pacific Islander; SOR, Some Other Race. The Hispanic population can be of any race and is not included in the race percentages, which add up to 100%.

SOURCES: Column (1): Monthly Postcensal Resident Population table for January 1, 2020, to June 1, 2020, at https://www.census.gov/programs-surveys/popest/technical-documentation/research/evaluation-estimates/2020-evaluation-estimates/2010s-national-detail.html. Column (2): Table NC-T2021-SR11H, Annual Estimates of the Resident Population by Sex, Race, and Hispanic Origin for the United States: April 1, 2020 to July 1, 2021, at https://www.census.gov/data/tables/time-series/demo/popest/2020s-national-detail.html. Column (3): See Table 10.3. Column (4): Estimates for race groups developed by Citro (2021).

(3) the 2020 Census as reported (with SOR); and (4) the 2020 Census with SOR Alone reallocated to specific race groups, by applying the percentage allocations of SOR Alone to specific race groups used in 2010 to 2020 SOR counts. The blended base population estimates for 2020 (column 2) closely resemble the original 2020 estimates carried forward from 2010 (column 1). They are dramatically different from the 2020 Census counts (column 3). This outcome is seen to some extent in every census given that the census has an SOR category and the population estimates do not; for 2020, the differences are exaggerated because of the new, expanded coding rules for the 2020 Census, which generated many fewer people in the White Alone category and many more in the Two or More Races category. Using a rough allocation method based on the 2010 Census, the 2020 Census estimates with SOR reallocated (column 4) are in between the 2020 blended base population estimates and the 2020 Census as reported counts (columns 2 and 3, respectively).

Copyright National Academy of Sciences. All rights reserved.

10.5.3 Noise in the 2020 Redistricting File

The Census Bureau warned that block data in the Redistricting File would need to be aggregated because of the new method for injecting noise into the data to protect confidentiality (see Chapter 11). The Census Bureau estimated, however, that once a block group had between 450 and 499 or more people, the ratio of the population represented by the largest race/ethnicity group from the publicly released Redistricting File would be within ±5 percentage points of the actual census counts at least 95% of the time. The same statement would apply to minor civil divisions and places with between 200 and 249 or more people. This assessment was based on comparing the 2010 Redistricting File (that used swapping for confidentiality protection) with a "demonstration" 2010 Redistricting File that used the production parameters for the 2020 Census confidentiality-protection method (Wright and Irimata, 2021:1). Nonetheless, the 2020 redistricting data are quite noisy for some types of areas and population groups.

Figure 10.5 graphs the percentages of counties, incorporated places, and census tracts with differences of 10% or more for race and Hispanic population groups between the 2010 Redistricting File protected with the 2020 DAS and the original 2010 file protected with a technique called data swapping. With the exception of the White Alone population and, for counties, the Hispanic population, more than 10% to more than 60% of units have differences of at least 10% or more. Clearly, these data are quite noisy, particularly for small race groups, such as NHOPI, and generally for incorporated places.[23] That the nation's 19,483 incorporated places do not always nest neatly within counties and are smaller in population on average than even census tracts (the median population in 2020 for incorporated places was 1,129, compared with 3,775 for census tracts) makes it difficult to inject noise and retain accuracy for incorporated places. Small population sizes also make it difficult to square accuracy and confidentiality protection for American Indian reservations and Alaska Native villages, with median populations of 400–500 people in 2020.[24] Yet incorporated places and AIAN jurisdictions are communities with governmental functions that have historically depended on accurate data from the decennial census.

[23]Final 2010 Demonstration File Detailed Summary Metrics file at https://www2.census.gov/programs-surveys/decennial/2020/program-management/data-product-planning/2010-demonstration-data-products/01-Redistricting_File--PL_94-171/2021-06-08_ppmf_Production_Settings/2021-06-08-data-metrics-tables_production-settings.xlsx, using the June 8, 2021, production settings for the application of the TopDown Algorithm to the 2020 Census Redistricting File.

[24]See Census Geographies Project (2022) for population size data. See also Liebler (2022).

Copyright National Academy of Sciences. All rights reserved.

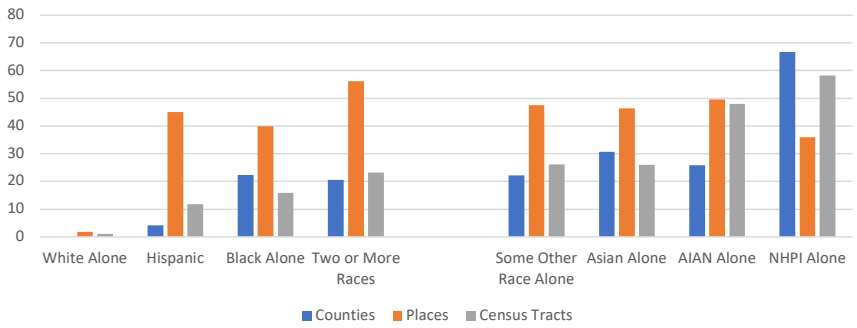


Figure 10.5 Percentage of counties, incorporated places, and census tracts with differences of 10% or more between the 2020 Disclosure Avoidance System privacy-protected and original 2010 redistricting files in estimates for race and Hispanic populations.

NOTES: AIAN, American Indian or Alaska Native; NHOPI, Native Hawaiian or Other Pacific Islander. Groups are ordered from left to right by their percent of total U.S. population.

SOURCE: Final 2010 Demonstration File Detailed Summary Metrics file at https://www2.census.gov/programs-surveys/decennial/2020/program-management/data-product-planning/2010-demonstration-data-products/01-Redistricting_File--PL_94-171/2021-06-08_ppmf_Production_Settings/2021-06-08-data-metrics-tables_production-settings.xlsx, using the June 8, 2021, production settings for the application of the TopDown differentially private algorithm to the 2020 Census Redistricting File.

10.6 RACE AND ETHNICITY MEASUREMENT IN 2020—CONCLUSIONS

Conclusion 10.1: The 2020 Census depicted a more diverse nation than did the 2010 Census. The White Alone population declined significantly, particularly among Hispanic people; the Hispanic population increased; and the Two or More Races population increased substantially, particularly among Hispanic people. It is unknown the extent to which these changes were due to demographic changes (e.g., more multiracial children); changes in self-identification; the provision of write-in origins (perhaps not intended to indicate racial identification) made possible by the format and processing changes to the 2020 Census race question; and other effects of the 2020 Census format and processing changes.

Conclusion 10.2: The 2020 Census had poorer quality data on race and ethnicity compared with the 2010 Census in terms of coverage error (net

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

overcounts and undercounts), rates of missing and imputed responses, and the noise infused by the 2020 Census Disclosure Avoidance System.

- *The 2020 Census exhibited larger coverage errors compared with 2010 for race and ethnic groups. There were increases in net overcounts for some groups, particularly non-Hispanic White Alone people and Asian people, and increases in net undercounts for other groups, particularly Black people, Hispanic people, and American Indians and Alaska Natives. Consequently, differences among groups widened significantly in 2020, with adverse implications for uses of census data to allocate fixed resources, such as representation, funding, and services.*
- *The 2020 Census had higher imputation rates (to account for missing and inconsistent responses) for race and ethnicity compared with 2010, particularly for responses not provided by a household member via the internet, a mailed-back paper questionnaire, or telephone. Moreover, the 2020 Census had much higher imputation rates for race and ethnicity for people in group quarters compared with the already high rates in 2010.*
- *The 2020 Disclosure Avoidance System injected noise of considerable magnitude into the Redistricting File for small race and ethnicity groups in small governmental jurisdictions, such as incorporated places and American Indian tribal lands and Alaska Native villages.*

10.6.1 Improvements to Collection of Race and Ethnicity Information

Our suggestions for improving of the collection of race and ethnicity data in the 2030 Census and beyond follow two basic tracks: first, continuing and extending lines of research on the decennial census and the ACS, and second, addressing the implications of possible revisions in federal standards for collecting race and ethnicity data.

Understanding how the changes in format and processes for the race and ethnicity questions in the 2020 Census and the ACS beginning in 2020 affected distributions may well be moot should the Census Bureau's preferred single combined race/origin question be approved for the 2030 Census. Yet, research on the matter is important for users to understand the implications for time series and for the Census Bureau to understand how respondents think of the responses they provide—for example, as a current self-identification of race or ethnicity or as an identification of long-past origins.

Types of research that could be valuable for users and the Census Bureau to understand the effects of the 2020 changes in format, data capture, and coding on race and ethnicity reporting include:

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

- Qualitative research with respondents varying in age, sex, race, and ethnicity. Research could include focus groups and one-on-one cognitive interviewing to ascertain how people viewed and used major category check boxes and write-in spaces.
- Tabular analysis of age, sex, race, and ethnicity reporting, comparing the 2010 Census, 2020 Census, the ACS before and after 2020, and the 2015 National Content Test to illuminate any trends, such as in multirace reporting by generations (e.g., baby boom, millennial).
- Simulation of the use of 2010 race and ethnicity questions and response processing for people in the 2020 Census, using a sufficient sample of age, sex, race, and ethnicity groups for efficient and statistically significant estimation.
- Simulation of the implications for the race distribution in states, counties, and other geographies of the higher frequency of write-in responses via the internet compared with other response modes.

Recommendation 10.1: The U.S. Census Bureau should conduct research to determine how the changes in format and processes that were made in the 2020 Census and in the American Community Survey beginning in 2020 affected the distributions of race and ethnicity. Such research should use qualitative, quantitative, and simulation methods to ascertain: how respondents viewed and used the 2010 and 2020 formats; trends in multirace reporting by age, sex, race, and ethnicity; how samples of 2020 respondents would have been categorized using the 2010 format, data capture, and coding; and the implications of differences in write-ins by response mode (e.g., more write-ins for internet responses) on race distributions among geographic areas. The Census Bureau should communicate the results of this research to data users to assist them to understand the implications of the changes for time series, and should also use the findings to inform 2030 Census planning.

Should OMB revise Statistical Policy Directive No. 15 to adopt a single, combined, check-all-that-apply question for race and ethnicity, with a MENA category added and Hispanic as a category, what might be the effects on race and ethnicity distributions in the 2030 Census? Table 10.7 gives a possible answer—a combined question, with write-in spaces under the White and Black checkboxes as in 2020 and the other 2020 data-capture and coding changes, would likely be a marked change in series from 2020, just as 2020 was a marked change in series from 2010. The SOR category would dwindle with a combined question, while the increase in the Two or More Races category and the decrease in the White Alone category would be somewhat less pronounced. Moreover, a combined question would likely show many fewer Hispanic people in the Two or More

Copyright National Academy of Sciences. All rights reserved.

Table 10.7 Percentage Distribution of the Population by Race/Ethnicity, 2010–2020 Censuses and Combined Race-Ethnicity Question

Source/ Race-Ethnicity Group	2010 Census	2020 Census	Combined Race-Ethnicity Question
Total (%)	100.0	100.0	98.8
White Alone	72.4	61.6	63.0
Black Alone	12.6	12.4	9.6
AIAN Alone	0.9	1.1	0.4
Asian Alone	4.8	6.0	5.2
MENA Alone	N.A.	N.A.	0.2
NHOPI Alone	0.2	0.2	0.1
SOR Alone	6.2	8.4	0.3
Two or More Races	2.9	10.2	9.3
Hispanic	*16.3*	*18.7*	10.8

NOTES: AIAN, American Indian or Alaska Native; MENA, Middle Eastern or North African; NHOPI, Native Hawaiian or Other Pacific Islander; SOR, Some Other Race. The Hispanic population in the 2010–2020 Census columns can be of any race and is not included in the race percentages, which add up to 100%; the Hispanic population in the combined question column is Hispanic Alone. The last column adds up to 98.8%—the remaining 1.2% were invalid or missing responses in the 2015 National Content Test (see Sources). The comparable estimate in the last column to the Hispanic estimates of 16.3–18.7% in the other two columns is 14% for Hispanic Alone and in Combination.

SOURCES: 2010–2020 Census: See Table 10.4. Combined Race-Origin Question: Mathews et al. (2017:Table H7 (all interview modes, row for a combined question with write-in areas)).

Races category (less than 5% in the 2015 National Content Test),[25] compared with one-third of Hispanic people who were so coded in 2020.

Other implications of the move to a combined question would need attention. First, the January 2023 *Federal Register* notice's query for feedback on requiring detailed/write-in race and Hispanic-origin data collection by default, whenever feasible, raises a very important issue. Many federal, state, and local administrative agencies (e.g., school districts, police departments, health departments, vital records offices) may not want or simply may not be able to use the detailed write-ins and additional codes that the Census Bureau has adopted. Instead, they may need or want to use the minimum, main categories only (White, Black, AIAN, Asian, Native Hawaiian and Pacific Islander,[26] Hispanic, MENA) to simplify data collection and processing. Even at the minimum category level, it could take considerable time for agencies to switch over to a combined question with Hispanic and MENA categories (e.g., all

[25]See Mathews et al. (2017:Table H13).

[26]The suggested changes in the *Federal Register* notice also include dropping "Other" from the NHOPI category (88 FR 5375, January 27, 2023).

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

states would have to adopt the new categories for vital records on births and deaths). Consequently, there will likely be periods when numerators (e.g., race and ethnicity of enrolled school children or newborns) do not correspond exactly to denominator populations derived from the census or ACS. For this reason, crosswalks would be essential for bridging series from the old to the new standards, as called for in the initial proposals for revision.

It would also be important to consult as soon as possible with state legislators, redistricting commissions, political parties, consultants, and others involved in the redistricting process. In this consultation, it would be useful to focus on ways to streamline the number of categories compared with 2000–2020. A combined question could increase the number of cells in Tables P1 (race or ethnicity) and P2 (race or ethnicity for 18 and older) because of the two added categories of Hispanic and MENA, but it would eliminate the need for two tables in the current list (P2, Hispanic and non-Hispanic by race, and P4, Hispanic and non-Hispanic by race for 18 and older—refer to Box 10.2). SOR could be ignored except for a single cell of SOR Alone, and redistricters might agree that stopping with, say, four or more races instead of six (or eight under the single-question format) would be acceptable. As part of this conversation, the issue could be raised about whether small blocks (the median block population size in 2020 was 28 people) could be combined. Streamlining and combining would decidedly help to square privacy protection with accuracy, whatever protection method were used.

Recommendation 10.2: To improve the quality of reporting on race and ethnicity in the census, American Community Survey (ACS), and other federal data collections, the U.S. Office of Management and Budget (OMB) should revise Statistical Policy Directive No. 15 to adopt a combined, check-more-than-one race/ethnicity classification with both Hispanic and Middle Eastern or North African as main categories, in addition to White, Black, American Indian and Alaska Native, Asian, and Native Hawaiian and Other Pacific Islander. OMB should formalize any changes to the race and ethnicity reporting standards as soon as practicable, to permit the U.S. Census Bureau to implement the new race/ethnicity question in the ACS and in the 2030 Census. OMB should allow agencies to collect data only for the major categories in the combined question when added detail (as in the 2020 Census) would impose undue administrative burdens. A decision by the Census Bureau on including added detail and write-in spaces, along with expanded data capture and coding procedures, for a combined race/ethnicity question for the ACS and the 2030 Census should await completion of research on such matters

Copyright National Academy of Sciences. All rights reserved.

as how respondents viewed the 2020 question format and the effects of response-mode differences in write-in responses on diversity among geographic areas.

Recommendation 10.3: The U.S. Census Bureau should produce a crosswalk or bridge between the 2010 and 2020 Census race and ethnicity questions and responses. Similarly, if a combined race/ethnicity question (with Hispanic and Middle Eastern or North African categories) is adopted as the standard by the U.S. Office of Management and Budget, the Census Bureau should produce a crosswalk or bridge between the 2020 version and revised race and ethnicity questions and responses, as soon as the revisions are implemented in the American Community Survey and then in the 2030 Census. The Census Bureau should involve data users in this important work.

Recommendation 10.4: The U.S. Census Bureau should consult as early as possible with the redistricting community, which includes state legislators, redistricting commissions, political parties, and political consultants, among others, to determine the optimum set of tabulations to include on the 2030 Redistricting File, whether the current race and ethnicity categories are retained or the proposed revisions take effect. The consultation should include consideration of streamlining the number of race/ethnicity categories in the file and combining blocks with small populations (using input from localities) to maximize the ability to protect confidentiality while maintaining accuracy.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

– 11 –

Impact of New Confidentiality-Protection Methods on 2020 Census Data Products

The panel's charge was to study the quality of the 2020 Census broadly defined (see Chapter 1) and make recommendations for 2030. Other chapters address the quality of 2020 Census data through the steps to produce the Census Edited File (CEF)—from generating the address list to enumerating residents to capturing and editing their responses, coding write-in answers to the race and ethnicity questions, and imputing values for missing answers. In this chapter, we assess the impact on data quality of the final step in creating tables and other products from the CEF for public use—namely, implementing one or more approaches to protect the confidentiality of the information.

Over the decades, the U.S. Census Bureau has used various methods (e.g., data reduction, cell suppression, swapping, noise injection using formally private algorithms—see Box 11.1 and McKenna, 2018) to reduce the chances that someone could "reidentify" a respondent from published census tables in reports and computer summary files.[1] The Census Bureau calls the set of confidentiality-protection methods used for a specific census or survey a Disclosure Avoidance System or DAS. The application of a DAS to data products is essential to honor confidentiality pledges to respondents and

[1]"Reidentification" involves "reconstructing" micro records (finest grain, individual person- or household-specific information) from published tables, followed by linking the records with outside sources to find potential matches—see Muralidhar and Domingo-Ferrer, 2023a).

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

preclude misuse of respondent information for administrative or enforcement (i.e., nonstatistical) purposes. By definition, however, disclosure avoidance impairs data quality relative to the collected data. The degree of impairment depends on the disclosure avoidance choices of the data producer. The question is how well the 2020 Census DAS preserved data utility while protecting confidentiality.

As described in Chapter 1, the decennial census is a national resource and the only reliable source of data for the nation's many small governmental jurisdictions. The return on the nation's sizeable investment in the census data collection comes from data products that are available on a timely basis and address users' needs for accurate data for small geographic areas and population groups, defined by race, ethnicity, and other characteristics. In this context, confidentiality protection is an important but not overriding objective of data product design and implementation, to the exclusion of the public's needs for useful data.

This chapter reviews the confidentiality-protection methods for the 2020 Census, which include traditional methods such as reduction of table content in most 2020 data products and cell suppression for detailed race and ethnicity tables (termed "adaptive design" by the Census Bureau for the particular form of cell suppression used in 2020). More importantly, the Census Bureau adopted algorithms that satisfy the concept of "formal" or "differential" privacy, developed by computer scientists in the mid-2000s, as the principal protection method for 2020, in place of the data-swapping methods used in the 1990–2010 Censuses.

We first provide an overview of 2020 Census data products that have been released or are still under development. Then we review the Census Bureau's decision to adopt formally private confidentiality-protection methods for the 2020 data products, describe challenges the Census Bureau encountered in implementation, and draw lessons for 2030. Throughout, we support the intention of the 2020 DAS, namely, to provide added protection in the face of threats to confidentiality from the readily available, sophisticated matching technologies and multiplicity of data on the internet. We find, however, that implementation of the 2020 DAS substantially delayed data release, resulted in sizable cutbacks in available data, and impaired the quality of some data (see Box 11.2).

Three appendixes support this chapter. Appendix F.1 provides a mathematical explanation of differential privacy. Appendix F.2 presents two important reference tables, one comparing the major 2020 Census and 2010 Census data products and the other describing the "demonstration files" generated by the Census Bureau, applying various settings of the 2020 confidentiality-protection methodology to 2010 Census data to enable feedback from data users. Appendix F.3 provides examples of the effects of the 2020 DAS on data quality for a range of uses.

Copyright National Academy of Sciences. All rights reserved.

Box 11.1 Methods for Protecting Confidentiality, 1970–2020 Census Summary Files

Data Reduction or Elimination is always part of confidentiality protection for census summary files, since publishing every tabulation for every geographic area would increase disclosure risks and make a file too unwieldy to use. Decisions on which tables to include offer the opportunity to reduce disclosure risks while satisfying important user needs. Forms of data elimination are used to provide tabulations for geographic areas with large populations (e.g., states and counties), but not areas with small populations (e.g., blocks and census tracts), and to reduce the content for small areas compared to larger areas (e.g., 5-year age groups vs. single years of age). Similarly, data can be eliminated or reduced for small racial or ethnic groups.

Cell Suppression is a targeted form of data elimination used for confidentiality protection for summary file tabulations in the 1970 and 1980 Censuses and again in 2020 (for the 2020 Detailed Housing and Demographic Characteristics Files—see Appendix F.2). The idea is to eliminate table cells with small values and not an entire table. Specifically:

- One or more cells of a table (e.g., Hispanic men ages 20–24) can be suppressed (deleted) if the value falls below a specified threshold (say, five people in an area).
- One or more additional cells of a table can be suppressed if the user could determine the value for a suppressed cell through subtraction (complementary suppression). For example, given six Hispanic people ages 20–24 in an area, of whom five are women, but there is an empty cell for men, there must be one Hispanic man in the 20–24 age range in that area. In that case, both the male and female cells would be suppressed and only the total number of Hispanic people ages 20–24 provided for the area.

Swapping was used in the 1990–2010 Censuses instead of cell and complementary cell suppression, which users disliked because of the holes in the data. In swapping, a small number of households have their characteristics interchanged—for example, a five-member household headed by a Black person in block A might be exchanged for a five-member household headed by a White person in block B. Swapping did not change the total number of people or households in a geographic area. Swapping targeted households deemed most at risk of disclosure, such as a Black household in a largely Asian or White area. The exact specifications were not made public by design to prevent reverse engineering.

Noise Infusion with Differential Privacy-Based Algorithms, adopted for 2020 Census data products, functions by applying controlled random noise to every cell in every table. For example, a table of sex by age by Hispanic origin with one Hispanic man and five Hispanic women ages 20–24 before the injection of noise might become two Hispanic men and six Hispanic women ages 20–24 afterwards. The amount of noise is controlled by the share of the total "privacy loss budget" (commonly denoted by ϵ, epsilon) allocated to a particular table. A series of post-processing steps then eliminate negative values, enforce consistency up and down the geographic hierarchy (e.g., census block totals add to block group totals to census tract totals), and address other anomalies of the noise injection in each product. The privacy budget parameters and the code used to implement the chosen algorithm can be made public. See Box 11.4 for the steps of the specific algorithm used to create the 2020 Redistricting and Demographic and Housing Characteristics Files.

Copyright National Academy of Sciences. All rights reserved.

Box 11.2 Effects of the 2020 Disclosure Avoidance System on Data Quality: Overview

The use of differentially private (DP) algorithms to protect the confidentiality of tables in 2020 Census data products through injecting noise into almost every cell was detrimental to data quality. This process:

- Delayed the release of data products, some of which are not scheduled for release until September 2024—in contrast, data delivery from the 1990–2010 Censuses was essentially complete by December of the year ending in 2 (e.g., 2012—see Table 11.1);
- Resulted in cutbacks in available data (see Table F.1 in Appendix F.2)—for example, there are many fewer tables that join characteristics of households and people (e.g., ages of children by household type), most of which will not be available for substate areas; and
- Reduced the quality and credibility of important data—for example, data show blocks with occupied housing and no people and vice versa, while tables for rural areas, renters, small towns and places, and small population groups (e.g., American Indians and Alaska Natives on reservations) are generally less accurate than tables for urban areas, owners, and large population groups (see Appendix F.3).

The DP approach offered theoretical advantages compared with data swapping:

- Enhanced protection against attempts to reidentify individuals from the tabular data by attackers with unlimited computing power and time and access to voluminous data sources to use in reidentification;
- Transparency regarding the tradeoffs between accuracy and confidentiality protection (captured in the value of the tuning parameters); and
- Ability to fully document the operation of the algorithm, including providing software code.

In practice, DP was ill suited to the requirements of census data users, which include:

- Timely information—but with DP, each data product had to be processed separately for confidentiality protection, which took more time than implementing protection once in the Census Edited File, as in swapping;
- Population counts for geographic areas unaltered by an algorithm, since census totals act as official numbers for many state and local government data uses—but adding unaltered population counts and other "invariants" compromises the ability of an algorithm to deliver a stated privacy guarantee;
- Numbers that are not negative and that consistently add up (e.g., the number of Black people in the blocks of a city should add to the city total);[a]
- Data that are equally accurate for geographic areas (e.g., census tracts) that "nest" within counties and states, and geographic areas (e.g., incorporated places and American Indian tribal areas) that often cross boundaries—but these requirements make it difficult for algorithms to efficiently process data or achieve the result of equal accuracy; and
- Data that link households and their members, such as persons per occupied housing unit or three-generational families—but these requirements are difficult for differentially private algorithms in their current state of development to handle without greatly reducing accuracy or protection.

[a]This text was revised after release of the prepublication version of the report to strike previous language about these conditions explicitly compromising privacy guarantees; instead, the non-negativity and hierarchical additivity complicate the interpretation of privacy guarantees.

Copyright National Academy of Sciences. All rights reserved.

Box 11.2 (continued)

For the first two major data products—the 2020 Redistricting File and the Demographic and Housing Characteristics (DHC) File—the Census Bureau employed a TopDown Algorithm (TDA) to construct a synthetic Microdata Detail File for use in tabulation instead of the Census Edited File. The TDA, a series of postprocessing steps based on the noise-infused tables or queries—functioned to:

- Eliminate negative and decimal values;
- Ensure additive consistency by geography, with lower-level geographies consistent with higher-level geographies (e.g., blocks add up to tracts, tracts to counties, counties to states, states to nation);
- Create special block groups to provide more accurate data for geographic areas that did not nest; and
- Ensure that some invariant constraints preserved values as-is (i.e., state population totals and block housing unit- and group quarters-type totals).

In addition, the Census Bureau set values of the tuning parameter ϵ much higher than believed protective in past applications of DP noise infusion, which, together with the TDA, produced accurate data for larger geographic areas and groups.

For the remaining major data products—Supplemental-DHC and Detailed Demographic and Housing Characteristics (DDHC) Files A and B—the Census Bureau contracted for the development of specialized DP algorithms (PHSafe for Supplemental-DC and SafeTab-P and SafeTab-H for the DDHC A and B Files). These algorithms are not all fully developed, and PHSafe can only produce accurate data for person-household variables by eliminating substate geographic detail.

In sum, the costs of the 2020 DAS were long delays to obtain a balance acceptable to the Census Bureau between accuracy and confidentiality protection (accounting for user feedback); undermining the privacy guarantee; inability to provide many linked household and person characteristics that users value; and variations in accuracy among geographic areas and population groups. Finally, users lack clear guidance and measures of error from the noise infusion and so find it hard to evaluate fitness of the data for their uses.

11.1 CENSUS DATA PRODUCTS: OVERVIEW OF CONTENT AND TIMING OF RELEASE

Historically, data from the decennial census were tabulated and made available in the form of printed reports. Publication of the reports could take years after data collection, but it was common practice to issue "preliminary" reports of total population very quickly after completion of data collection and "advance" reports with specific tables that were later bound into books. Although only a fraction of the data collected could be printed in the precomputer age, small-area data were included in census reports from the beginning: for counties and minor civil divisions back to 1790; for incorporated places back to 1880 (and even further back for some large cities); for census tracts in the cities that pioneered these neighborhood areas back to 1910;

Copyright National Academy of Sciences. All rights reserved.

for enumeration districts (the equivalent of today's block groups) back to 1930; and for blocks in large cities back to 1940 (housing characteristics) and 1960 (population and housing characteristics) (Census Geographies Project, 2022:Ch. 3).

The 1960 Census saw a major expansion of geographic and subject detail included in printed reports. In addition, Summary Files of aggregate statistics (e.g., a table of single years of age by race and ethnicity for census tracts) and Public Use Microdata Samples (PUMS) were produced at the specific request of and with funding from outside groups (U.S. Census Bureau, 1966). Computer data products became standard beginning in the 1970 Census.

Four major computer summary products from the 1990–2020 Censuses are listed below (see Table F.1 in Appendix F.2 for detailed information on their content and other features for 2010–2020).[2] They are:

- 50 state total population counts (including federal employees overseas) for reapportionment of the U.S. House of Representatives;
- Redistricting File of race and Hispanic origin tables for the total population and voting-age population (ages 18 and older) for all census geographies, including but not limited to states, counties, towns, places, school districts, American Indian tribal areas, Alaska Native villages, census tracts, block groups, and blocks;
- Summary File 1 (SF1), renamed for 2020 as the Demographic and Housing Characteristics (DHC) File—a workhorse file for federal, state, and local governments and other users, with such information as single years of age by race and ethnicity, household type, and owner/renter status for all census geographies down to the block level (in 2020, a Supplemental-DHC File is to provide a limited set of tables of linked household and person characteristics with greatly curtailed geographic detail); and
- Summary File 2 (SF2), renamed for 2020 as the Detailed Demographic and Housing Characteristics (DDHC) File with two subfiles A and B, which provide data for hundreds of race and ethnicity categories on age and sex (DDHC-A) and household type and housing tenure (DDHC-B) with less geographic detail than SF1/DHC.

The 1990–2010 Censuses had processes in place to generate data products on a timely basis; see Table 11.1. The 50 state numbers for reapportionment

[2]Each census has released additional specialized tabular products, such as Demographic Profiles and Congressional District Files, and PUMS files—these files are not further discussed except that Table F.1 summarizes content and release schedules for the 2010 and 2020 Demographic Profile and PUMS files. In addition, the 1970–2000 Censuses released data products from large samples of the population that received a "long-form" questionnaire, which obtained data on education, income, housing costs, and many other variables. These variables are now collected in the American Community Survey.

Copyright National Academy of Sciences. All rights reserved.

Table 11.1 Release Dates/Scheduled Dates for Four Census Data Products, 1990–2020 Censuses

Census/Year/ Data Product	Release Date			
	Census Year	Census Year + 1	Census Year + 2	Census Year + 3
1990				
Reapportionment	Dec			
Redistricting		Jan–Mar		
SF1/DHC		Mar–Jun		
SF2/DDHC		Sep–Dec		
2000				
Reapportionment	Dec			
Redistricting		Mar		
SF1/DHC		Jun–Aug		
SF2/DDHC			Jan–Apr	
2010				
Reapportionment	Dec			
Redistricting		Feb–Mar		
SF1/DHC		Jun–Aug		
SF2/DDHC		Dec–	-Apr	
2020				
Reapportionment		April		
Redistricting		Aug		
SF1/DHC				May[a]
SF2/DDHC				Sep[b]

[a] DHC was released on May 25, 2023; Supplemental DHC is scheduled for September 2024 release.

[b] DDHC-A was released on September 21, 2023; DDHC-B is scheduled for September 2024 release.

NOTES: SF, Summary File; DHC, Demographic and Housing Characteristics (File); DDHC, Detailed DHC (File). Range of months indicates product was released state by state. In 2020, reapportionment and redistricting data were delayed because of delays in enumeration and processing from the COVID-19 pandemic. In 2000 and 2010, a national file for SF2 was released in December of Census Year + 2. SF1 and SF2 were delayed in the 1980 Census (not shown) because of budget cuts in the first term of the Reagan administration.

SOURCE: Compiled by panel from Census Bureau public releases.

Copyright National Academy of Sciences. All rights reserved.

were released 9 months after Census Day (April 1), and a Redistricting File (replacing the old preliminary and advance reports) was released 12 months after Census Day as provided by statute (13 U.S.C. § 141(b)–(c)).[3] SF1 (now DHC) was typically released state-by-state 2–5 months after the Redistricting File. SF2 (now DDHC-A, B) was typically released 5–8 months after SF1. Data release was essentially complete 2 years after Census Day.

In contrast, 2020 Census products' preparation and release processes have been far from smooth or timely. Delays in the enumeration consequent to the COVID-19 pandemic led to delays in release of the reapportionment state counts (released April 2021) and the Redistricting File (released August 2021). The 2020 DHC File was released on May 25, 2023, and the DDHC-A File was released on September 21, 2023. The DDHC-B and Supplemental-DHC Files are not scheduled for release until September 2024. The delays in releasing the DHC and subsequent products are primarily due to the Census Bureau's decision to adopt formal privacy methods at a late stage in census planning. The new approach had not previously been tested in the census context, and the Census Bureau had no backup plan when implementation proved more difficult and time consuming than expected.

With regard to content, the 2010 and 2020 Redistricting Files are virtually the same in terms of meeting the needs of the redistricting community. The 2020 Redistricting File added a table of people in group quarters by type (e.g., nursing home, dormitory, etc.).

The 2020 DHC File has reduced content compared with the 2010 SF1.[4] For example, the 2020 DHC File has 36 tables for blocks as the lowest level of geography compared with 73 block-level tables in the 2010 SF1 File; similarly, the 2020 DHC File has 19 tables for census tracts as the lowest geographic level compared with 28 tract-level tables in the 2010 SF1 File. Some amount of table deletion is sensible for added confidentiality protection, although only a comprehensive user survey could determine which deletions were broadly acceptable and which caused problems for long-established data uses. Some deletions are known to be problematic. For example, item nonresponse indicators were provided in 2010 for blocks, which users found helpful in deciding which areas to flag as likely having lesser quality data; but these indicators were deleted altogether in 2020, and tables that link household or family and person characteristics were cut back for all geographies. Some of the deleted household or family and person "join" variables will be provided in the Supplemental-DHC File in September 2024—for the nation and states only, and not for any substate geographies as originally planned.

[3]The current form of the deadlines in 13 U.S.C. § 141(b)–(c) dates to amendments in 1975–1976.

[4]On the other hand, the Census Bureau did add some tables to the 2020 DHC that were not previously available, including sex-by-age tables for all combinations of race/ethnicity.

Copyright National Academy of Sciences. All rights reserved.

The 2020 DDHC-A File and forthcoming DDHC-B File have greatly reduced content compared with the 2010 SF2 (and a companion American Indian or Alaska Native File). DDHC-A has two tables for race and ethnicity groups with at least 22 people in a census tract or other area: (1) total population; and (2) sex by selected age categories depending on the size of the group. Similarly, DDHC-B has two tables: (1) household type (number of categories depending on the size of the group); and (2) tenure (owner/renter). In contrast, SF2 in 2010 had 61 tables for census tracts as the lowest level of geography and 10 tables (for people in group quarters) for counties as the lowest geographic level.[5] Also, the DDHC-A and -B Files provide tables for a subset of types of geography including cities but not towns, although towns play an important role in local government in many states.

The 2020 DAS design and implementation process generated data-user concerns beyond time delays and reduced content and geographic detail. These concerns, expressed in a letter to the Census Bureau director (Federal State Cooperative on Population Estimates Steering Committee, 2022), focused on accuracy and utility and included:

- Anomalies in the 2020 Redistricting File and in "demonstration" files for the DHC File, which call into question the data's accuracy and utility, particularly for small geographic areas and population groups;[6]
- Lack of adequate guidance for how to interpret or explain the error introduced by the DAS;
- Concerns that the TopDown Algorithm used to protect the 2020 Redistricting and DHC Files cannot provide acceptable accuracy for variables that link or join person characteristics with housing unit characteristics;[7]
- Indicators of bias in the demonstration files for the 2020 Redistricting and DHC Files that give rise to concerns about data equity (e.g., for rural versus urban areas, more diverse versus less diverse areas, majority rental versus majority owner-occupied areas); and

[5]The 2010 SF2 had a higher population threshold (at least 100 people in an area) than the DDHC-A and -B Files but did not vary the number of categories for any tables.

[6]On inaccuracies in the Redistricting File, see Kenny et al. (2021) and Appendix F.3; also see Appendix F.3 for examples of anomalies in the DHC File based on using 2010 data to demonstrate the expected operation of the 2020 DAS. The 2020 DAS also impairs comparability with previous censuses for small geographic areas and population groups.

[7]The Census Bureau acknowledged this problem from the outset and warned users not to calculate persons per occupied housing unit from the Redistricting File. The Census Bureau included several join variables in the DHC File by adding recodes to the household head's record for input to tabulations (e.g., size of household, presence of people ages 65 and older, presence of own children under age 18). Additional complex join variables, but many fewer than in 2010 and none for substate geographies, are to be included in the Supplemental-DHC File, developed with a different algorithm—see Table F.1.

Copyright National Academy of Sciences. All rights reserved.

- Concerns that the delays in releasing DHC and DDHC data are affecting other Census Bureau programs, including population estimates and coverage evaluation via Demographic Analysis (DA).

11.2 CONFIDENTIALITY PROTECTION IN THE DECENNIAL CENSUS

The Census Bureau collects census and most survey data under a pledge of confidentiality,[8] which is meant to respect respondents, establish and maintain public trust, and foster the willingness of respondents to provide accurate data to the census and surveys. Indeed, since 1929, census law has stipulated that census employees must not release data that could "identify" an individual (similar protection for businesses dates to 1909—see Box 11.3). Census employees, along with external researchers, contractors, and others who have passed background checks and have Special Sworn Status, face substantial penalties for disclosure of individual information. The law does not specify what "identify" means—specifically, whether it pertains solely to such individual identifiers as name and address, or whether it also pertains to individual attributes (i.e., characteristics such as age, sex, race, household relationship), or whether it also pertains to group attributes (e.g., everyone in census block 101 is of White non-Hispanic origin). The Census Bureau has come to adopt a broad interpretation of the information protected under census law, including not just personally identifying information in census or survey responses but information about general attributes of respondents and operational paradata on the mechanics of response (e.g., whether a person self-responded or if a household did not respond).[9]

Census data products are never released in raw form but always include protections to minimize the risk of disclosure of an individual's responses.[10] As summarized in Box 11.1, the Census Bureau has used a wide variety of disclosure-protection methods, including table and cell suppression and, in the 1990, 2000, and 2010 Censuses, data swapping—that is, exchanging some household records among small areas to try to thwart direct one-to-one identification of census returns based on auxiliary data. The Census Bureau came to believe that swapping did not provide sufficient protection, although

[8]The exception is surveys of state and local governments, through which the Census Bureau collects publicly available information on employment and finances.

[9]See, for example, the material for annual data stewardship training for Census Bureau staff, the fiscal year 2023 version of which is at https://broadcast.census.gov/main/training/2023/fr/dscui/t13/story.html.

[10]The exceptions are that microdata records from the census are released by the National Archives 72 years after a census and that an individual (or their heir) may request their own records from censuses not yet released by the National Archives through the Census Bureau's Age Search service. See https://www.census.gov/history/www/genealogy/decennial_census_records/census_records_2.html.

Copyright National Academy of Sciences. All rights reserved.

Box 11.3 Legislation Protecting the Confidentiality of Decennial Census Information

Section 9 of Title 13 of the U.S. Code specifies:

> Neither the Secretary, nor any other officer or employee of the Department of Commerce or bureau or agency thereof, may, except as provided in section 8 of this title. . .
>
> (1) use the information furnished under the provisions of this title for any purpose other than the statistical purposes for which it is supplied; or
> (2) make any publication whereby the data furnished by any particular establishment or individual under this title can be identified; or
> (3) permit anyone other than the sworn officers and employees of the Department or bureau or agency thereof to examine the individual reports.

A form of this language dates to the 1909 Census Act authorizing the 1910 Census, which provided that "no publication shall be made by the Census Office whereby the data furnished by any particular establishment can be identified" (36 Stat. 9). Protection of personal information was added 20 years later when "individual" was added (46 Stat. 25). When Congress codified Title 13 in August 1954, this provision was renumbered but not substantially altered, and the core language has not been amended since.

13 U.S.C. § 214 provides for punishment of census employees and sworn agents who publish or communicate confidential data under Section 9 through a fine of not more than $5,000 or a prison term of not more than 5 years, or both. 44 U.S.C. § 3572, added by the Foundations for Evidence-Based Policymaking Act of 2018, increased the punishment (for employees or sworn agents of any federal agency collecting data under a pledge of confidentiality) to a fine of not more than $250,000 or a prison term of not more than 5 years, or both.

NOTE: The Second War Powers Act of 1942, Section 1402, which expired in March 1947, suspended Section 11 of the 1929 Census Act, thereby permitting census data to be used to identify individuals "in connection with the conduct of the war."

there has been no known reidentification of individuals outside the Census Bureau from swapped summary files. Swapping also introduced error, like any disclosure-protection method, but did not change the number of people and households in an area, and swapped data from previous censuses have passed users' face validity test for reasonable accuracy.

The Census Bureau decided to adopt a new DAS for the 2020 Census, premised on the cryptographic concept of differential privacy.[11] A differential

[11]The relevant computer science literature on formal privacy references "privacy" and not "confidentiality." Statistical agencies have traditionally distinguished the two, defining "privacy" as control of one's information, which they respect by asking only for information that is needed for public purposes and does not violate long-standing public concerns (e.g., U.S. statistical agencies do not ask about religion because of separation of church and state). Respondents give up control when they reply to censuses and surveys. In return, they are assured their information will be used for statistical purposes only and not against them as individuals (e.g., not shared with immigration or

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

privacy-based algorithm would infuse a controlled amount of random noise into all of the tabular cells of each data product, in the context of an overall privacy-loss budget allocated to various tabulations to reflect the tradeoff of disclosure risk and accuracy. The actual noise added to a specific cell count would remain unknown, but the distributional properties of that noise would both be known and publicly shareable. The privacy-loss budget is commonly denoted by ϵ (epsilon); even though later versions of the Census Bureau's algorithms are parameterized by rho and delta (ρ and δ), we use ϵ as shorthand for the parameters in this report because ϵ has been the basis for most discussion with data users and the public.[12] (See Appendix F.1 for a mathematical explanation.)

Any attempt to reduce disclosure risk necessarily compromises data utility. In contrast to data swapping, the new approach promised to be transparent and to explicitly acknowledge the tradeoff between data utility and confidentiality protection.[13] It also promised to respond to the threats to confidentiality protection from sophisticated matching technologies and the availability of voluminous data sources about people on the internet, which could be used to try to reidentify individual respondents to the census from published tables.

11.3 THE ROAD TO THE 2020 DISCLOSURE AVOIDANCE SYSTEM

Planning a decennial census is a decade-long (and often longer) enterprise, building on lessons learned from the previous count. The Census Bureau develops an "operational plan," initially published mid-decade, which sets forth its goals and challenges (see Chapter 2). As the decade progresses, it revises the plan and issues other public statements about decisions made and still to be made.

11.3.1 What Did the Operational Plan and Other Statements Say About Disclosure Avoidance?

The original 2015 operational plan design for disclosure avoidance simply stated that disclosure avoidance was a step in response processing in preparation for production of data products (U.S. Census Bureau, 2015:123). This statement remained as late as Version 3 of the operational plan (U.S. Census Bureau, 2017:128). Both statements located disclosure avoidance at the stage of

tax authorities). They are also assured that the agency will strive to protect their information against breaches of confidentiality—that is, disclosure of individual identities. The difference between the two terms—privacy and confidentiality—is subtle but important in framing the debate on data utility versus protection.

[12] This text was added after the prepublication release to clarify the report's usage of ϵ parameter notation due to its prominence in public discussion of disclosure avoidance in the 2020 Census.

[13] A recent paper by Bailie et al. (2023) places swapping into the differential privacy framework.

Copyright National Academy of Sciences. All rights reserved.

processing the CEF to produce a Hundred Percent Detail File (Version 1) or a Microdata Detail File (Version 3) as input to all the data products, as was done in 2010.

In the original plan, there was no indication that disclosure avoidance would require major changes to the delivery and content of data products. Nor had there been external pressure for changes in the disclosure avoidance method, akin to the complaints about costs, undercounts, and the quality of the Master Address File (MAF) or the calls for greater utilization of administrative records for data collection (see, e.g., National Research Council, 2004a, 2010, 2011).

Concurrently, in the global context of the emergence of Big Data and sophisticated linkage technologies, computer scientists in academe and the private sector were making technical advances in the theory and practice of curating and protecting large datasets. The work was designed to protect the identity of individuals (e.g., people posting on Facebook) with new methods that would allow for research on these rich, new data sources but not leak information that could potentially harm subjects (Herdağdelen et al., 2020).

To produce new data products from highly sensitive economic datasets, the Census Bureau researched and adopted the new protection methods for products such as OnTheMap.[14] Late in the 2010s, the Census Bureau's Research and Methodology Directorate proposed applying those new methods to the 2020 Census. The Census Bureau made a presentation to the Census Scientific Advisory Committee (CSAC) in fall 2016 that argued for a new approach for confidentiality protection; in a fall 2017 presentation to CSAC, the Census Bureau indicated it planned to use differential privacy-based algorithms to infuse noise into the 2020 Census data products (Abowd, 2016; Garfinkel, 2017).

During 2016–2018, the Census Bureau simulated an "attack" on 2010 data products using commercial and publicly available data. The simulation involved first "reconstructing" individual-level records from 2010 published tabulations, then matching the commercial data (which contained name, address, sex, and age) to the reconstructed individual-level records on age, sex, and census block but not race or ethnicity. The Census Bureau estimated that 17% of the U.S. population could have their race or ethnicity identified with a high level of certainty based on comparing the matched commercial-reconstructed records to the actual responses in the CEF (which, however, an outside attacker could not do).[15]

[14]OnTheMap (introduced in 2008) displays where employed people live and work at the block group level. It uses formally private synthetic data on the residence side, an early variant of differential privacy called probabilistic differential privacy. See Dajani et al. (2017:Slides 2, 18) and Abowd (2016).

[15]The Census Bureau reported a 58% confirmed match rate by using the CEF as the outside data source in place of the commercial data, on the assumption that outside data sources would improve in accuracy in the future (see Abowd, 2021a:Table 3). In subsequent work, the Census Bureau has

Copyright National Academy of Sciences. All rights reserved.

The description of this simulation is available only in abbreviated form in presentations and with somewhat more detail in affidavits filed in *State of Alabama v. U.S. Department of Commerce* (2021) (Abowd, 2021a,c). A full description is said to be in peer review for publication in a journal[16] but was not made available to this panel. There are criticisms of the simulation in the literature (e.g. Muralidhar and Domingo-Ferrer, 2023a),[17] but detailed criticisms are precluded by the absence of the complete methodology. Similarly, there is not sufficient information to assess the findings in Abowd and Hawes (2023:Table 5) of a lower but not negligible confirmed match rate for the 2020 DAS (noise infusion with the TopDown Algorithm, TDA) compared with the 2010 swapping and an "enhanced" swapping. (These comparisons matched the CEF as the outside data source to a reconstructed DHC demonstration file.) That information is to be in the as-yet-unpublished paper.

Census Bureau executives initially said that the results of the original reconstruction-reidentification study did not represent a major risk for the public in responding to the census.[18] At the same time, they said that the simulation provided strong justification for abandoning data swapping for a new disclosure-avoidance method based on formal privacy (e.g. Abowd, 2019b).

In July 2018, the Census Bureau issued a *Federal Register* notice calling for data users to explain and justify their uses of 2010 Census data tables, to finalize the roster of 2020 Census data products. After the fact, the notice's call for feedback could be characterized as an attempt to prioritize use cases, to allow the Census Bureau to precisely tune the noise that would eventually be added to 2020 Census tables, to ensure greater accuracy for higher-priority uses (at the necessary expense of lesser accuracy for other uses). But, in the moment, that purpose was not specified; instead, the notice included only the cryptic warning that "given the need for improved confidentiality protection, we may reduce the amount of detailed data we release to the public" (83 FR 34111, July 19, 2018). Respondents to the notice were unclear of the specific purpose, found it difficult to work with the Excel spreadsheet provided as a template for response, and expressed concern over where and to what extent product content might be cut back (Hotz and Salvo, 2022). Table F.1 indicates that substantial amounts of content were dropped or provided for higher-level geographies for both the 2020 Census DHC and DDHC products compared with the 2010 Census SF1 and SF2 products.

---

reported confirmed match rates of 90% and above by allowing age to vary by ±2–5 years for people older than age 21. See Abowd and Hawes (2023:Section 3).

[16]Personal communication from S. Keller, U.S. Census Bureau, to J. Salvo, March 9, 2023.

[17]Muralidhar and Domingo-Ferrer (2023a) is accompanied by a comment by Garfinkel (2023) and a rejoinder from the authors (Muralidhar and Domingo-Ferrer, 2023b).

[18]For example, Jarmin (2019a) stated in a February 2019 blog post that "the accuracy of the data our researchers obtained from this study is limited, and confirmation of re-identified responses requires access to confidential internal Census Bureau information."

Copyright National Academy of Sciences. All rights reserved.

Finally, in December 2018, Version 4 of the 2020 Census operational plan (U.S. Census Bureau, 2018:139) marked the shift to what would become the new DAS:

> *Changes Made Since Version 3.0 Operational Plan Release:* [The Data Products and Dissemination operation] now includes the application of the 2020 disclosure avoidance methodology to the microdata[19] in order to produce the Microdata Detail File. The process was previously part of the Response Processing Operation (RPO). The disclosure avoidance methodology that will be implemented for the 2020 Census is known as differential privacy. Differential privacy is the scientific term for a method that adds "statistical noise" to data tables we publish in a way that protects each respondent's identity.

The plan continued (U.S. Census Bureau, 2018:140) that "this new methodology will be tested and implemented with the 2018 End-to-End Census Test." Version 5 of the operational plan repeated the statement (U.S. Census Bureau, 2022b:145). Further public announcements followed in a February 2019 blog by deputy director Ron Jarmin and a presentation to the media in February 2019 by chief scientist John Abowd at a meeting of the American Association for the Advancement of Science. These statements engendered concern, questioning, and criticism from the data-user community (see, e.g., Ruggles et al., 2019).

11.3.2 Implementation Issues

What Versions 4 and 5 of the operational plan did not say, nor did the public announcements in early 2019 convey, was that implementation of the revised DAS was fraught with problems from the outset. Version 5 did note (U.S. Census Bureau, 2022b:145–146):

> To complete the implementation of the new disclosure methodology for 2020 Production, the entire 2020 Census data products suite must be defined in advance. The process for determining the 2020 Census data products began with the announcement of the *Federal Register* Notice to solicit data user feedback. The assessment phase is now underway with a final determination of 2020 Census products expected to be completed in early 2019. The Census Bureau anticipates publishing the plans for the 2020 Census data products in a future notice. . . .
>
> The 2018 and 2020 Disclosure Avoidance System (DAS) algorithms are highly complex, under active development, address subtle though precisely stated mathematical privacy issues, and solve genuinely novel outstanding scientific problems. All of these factors create an environment where only

[19] "Microdata" is a misstatement—the 2020 DAS would be applied to tabular data produced from the CEF, which would then be turned into synthetic microdata for input into the Census Bureau's tabulation software.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

rigorous software engineering standards, code review, and formal external auditing of internal code can be expected to reasonably remove major bugs from the code.

In Box 11.4, we summarize the basic steps in the TDA used in the production of the 2020 Census Redistricting and DHC Files. Tracing through the steps in the process, the complicated procedural issues and challenges (both anticipated and unanticipated) come into clearer relief, from determining the functional form of the noise infusion to allocating portions of the privacy-loss budget ϵ to subsets of variables and geographies. More importantly, the steps in the algorithm speak to the importance of three particularly fundamental disconnects between the methodology and its application in the census context to meet users' data needs. First, census data users need non-negative, integer numbers for official purposes, with at least some of these basic counts being held immune to any change (the "invariants" issue). Second, data users need variables that intersect person and housing characteristics, such as children in specific household types, and this is greatly complicated by the TDA's separation of person and housing data tables at the outset (the "joins" issue). Third, data users need variables and data for a plethora of political and statistical geographies, many of which are small in population or do not "nest" (the "off-spine" issue) or both. These three issues, as we discuss below, proved especially challenging for the approach, leading to manifold changes in the DAS implementation plans such that the statement in Version 5 of the operational plan (U.S. Census Bureau, 2022b:145) that "a final determination of 2020 Census products [is] expected to be completed in early 2019" proved aspirational.

The Problem of Invariants

Differential privacy theory for disclosure avoidance is designed to deploy noise to all cell counts in a set of cross-tabulations, to protect against current and future threats to confidentiality, without having to specify attackers' motives or estimate the plausibility of a particular attack. A requirement for some cell counts to remain "invariant," that is, reported as their actual collected values, makes it more difficult to inject noise in ways that do not markedly alter other cell counts or, indeed, undercut the theoretical basis of the privacy guarantee. Early on, the Census Bureau decided that the state population totals for congressional reapportionment would have to be invariant because of the constitutional language about "direct enumeration" of the population.

In a September 2017 presentation to CSAC, the Census Bureau further stated that much more would be held invariant—block-level population counts (not just state-level), voting age population counts, and occupied and vacant housing unit counts. These items were to be held invariant because they were also held invariant under confidentiality-protection methods used in the 2000 and 2010 Censuses, and because of a commitment made by the Census

Copyright National Academy of Sciences. All rights reserved.

Box 11.4 Basic Steps in TopDown Algorithm for 2020 Census Disclosure Avoidance

For the Redistricting File:

- *For person data:*
  - Define invariants (values to be held constant and reported as counted) and constraints to be applied (i.e., requirement of nonnegative integer totals); for the 2020 Redistricting File, total population at the state level (including the District of Columbia and Puerto Rico) was the sole person-level invariant.
  - Define the full set of queries (similar to tables) to be produced, effectively calculating the marginal totals for every combination of variable (8 levels of household/group quarters type $\times$ 2 levels of voting age $\times$ 2 levels of Hispanic origin $\times$ 63 levels of race) and geographic level (to the block level) from the Census Edited File (CEF).
  - Allocate shares of ϵ (or ρ) privacy-loss budget to query $\times$ geography combinations.
  - Add discrete Gaussian statistical noise to every cell in the queries, basing the magnitude of random noise on the ϵ (or ρ) share; resulting "noisy measurements" go forward in processing, and CEF's role is finished.
  - *At national level*, use optimization techniques to find best-fitting national table based on the nation-level noisy measurements, invariants, and constraints.
  - *At state level*, use optimization techniques to find best-fitting state-level tables based on the state-level noisy measurements, invariants, and constraints—adding the constraint that the state-level tables sum to the nation-level tables (technically, this work occurs at the state $\times$ American Indian/Alaska Native/Native Hawaiian [AIANNH] area levels, as AIANNH areas are processed in parallel to improve accuracy, and same holds true for subsequent levels of geography).
  - *At county level*, use optimization techniques to find best-fitting county-level tables based on the county-level noisy measurements, invariants, and constraints—now adding constraint that the county-level tables add to the state-level tables.
  - *Repeat optimization step—and incremental build-out of geographic detail—at census tract, block group, and census block levels.*
    - In these finer geographic levels, provision is made to consolidate ϵ shares for "parent" geographic units with only one "child" (i.e., a county containing one and only one tract) and then only using the parent's noisy measurement, to avoid multiple directly contradictory measurements applying to the same unit.
    - In the TopDown Algorithm, block groups are specially defined—different from census tabulation block groups—to better align with meaningful political or legal geography (i.e., creating "optimized" block groups that can be combined to form incorporated places in fewer pieces than if assembled from census blocks alone, and thus improving their accuracy).
  - Convert the final optimized block-level tables into person-level Privacy-Protected Microdata File (PPMF), which then carries forward into tabulation and data product generation.

Copyright National Academy of Sciences. All rights reserved.

Box 11.4 (continued)

- *For housing data:* Repeat all of the steps, albeit with:
  - Different variables: for the Redistricting File, only occupancy status (occupied/vacant) × geographic level.
  - Different invariants: count of housing units and count of occupied group quarters (by 7 major types), at the block level (and so at all higher levels of geography), but *not* the population count in those housing units or group quarters.
  - Convert final optimized block-level tables into housing-level PPMF and use for tabulation.

For the Demographic and Housing Characteristics Files: Repeat same process, with different (and more detailed) variables and combinations thereof—and adding constraints that summary totals be consistent with values in the Redistricting File.

Bureau to the U.S. Department of Justice for the 2000 Census, when statistical adjustment of the census counts for coverage error had been under consideration (Garfinkel, 2017). In a December 2018 presentation to CSAC, however, the Census Bureau proposed eliminating some of the invariants because of the difficulties they posed for privacy protection and implementation of noise infusion and the TDA (Abowd, 2018). The presentation previewed that the only invariants for the 2020 Redistricting File would be state-level population counts and block-level counts of housing units and the number and type of group quarters. The presentation acknowledged that the noise-infused population counts for blocks and other substate areas would need to add up to the invariant state totals. The reason for invariant total housing units and number and type of group quarters for blocks was because such counts are regularly available to localities working with the Census Bureau to update the MAF. Having the number and type of group quarters invariant would also ensure that a correctional facility, for example, would not appear in a block that had no group quarters or had a nursing home instead. Though the reduction in invariants was foreshadowed by Abowd (2018), the decision on invariants for the first set of 2020 Census data products was not finalized for another two years—in November 2020—and continued to generate user concerns and proposals for alternative solutions (Federal State Cooperative on Population Estimates Steering Committee, 2020).[20]

The actual extent to which the population and housing invariants undercut the privacy guarantee for the 2020 Redistricting and DHC Files is unclear. From the viewpoint of localities, the lack of substate invariant population totals

[20]The invariants were announced in a November 25, 2020, press release, reflecting a decision by the Census Bureau's Data Stewardship Executive Policy Committee the day before; see https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-updates/2020-11-25.html.

Copyright National Academy of Sciences. All rights reserved.

contradicted the idea of a census and raised concerns about the quality of the published data.[21]

The Problem of "Joins"

The census is a count of individuals within housing units, families, or living situations, generated by contacting a master file of addresses, identifying one person in the living unit as the individual (Person 1) around whom everyone else (Person 2, Person 3, etc.) is arrayed. The "household head," or "householder," is responsible for reporting on that household. In turn, data users have many uses for variables that link or join persons to their households, such as the number of persons per household (occupied housing unit).

Traditional census disclosure-control methods had no problem dealing with the hierarchical nature of the collected data. So, for example, in 2010, swapping was done by moving an entire household to a different geography, not by plucking out one person in a household and swapping them with someone in a different household. In contrast, the TDA used to produce the 2020 Redistricting and DHC Files infused (discrete) Gaussian noise into the person and housing data separately (the Gaussian distribution replacing the discrete Laplace noise infusion originally used—see Table F.2 and Appendix F.1). The TDA post-processing then continued separately for person and housing data, ultimately deriving synthetic microdata for persons and housing units with no direct linkage between the separate files.[22] Hence, the TDA had the desirable property of producing non-negative, integer numbers for each table cell that added up geographically (i.e., block populations added to census tract populations to county populations to state populations) but it could not make the person and housing tables consistent. Consequently, calculating persons per occupied household from the Redistricting File often gave anomalous answers (e.g., values of < 1 person per household), while some blocks had people but only vacant housing or only occupied housing but no people.

A different kind of differential privacy-based algorithm was needed to handle variables that required "joining" person and housing characteristics. The Census Bureau is developing the PHSafe algorithm for the join tables to be included in the Supplemental DHC File, but it did not have such an algorithm available at the time of producing the Redistricting File or the DHC File.[23] Yet person-household joins are of vital interest to state and local governments and other

[21]Personal communications from members of the Federal-State Cooperative Program for Population Estimates to J. Salvo.

[22]The reason for processing the person and housing tables separately is said to do with a technical issue—namely, that the sensitivity parameter in the noise infusion model could not be specified for households, due to their varying size. See Abowd et al. (2022:§ 5.3).

[23]As noted in an earlier footnote, the Census Bureau included some person-household variables in the DHC File through adding recodes to records for the household head on the CEF prior to creating tables for input to the noise infusion step.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

data users for projecting, locating, and evaluating services and programs; for validating their own intercensal estimates of population and housing; and for analyzing demographic and social trends.[24]

The Problem of Small Population, Off-Spine Geography

The United States has a large and bewildering variety of political and statistical units of geography, with considerable variation across states. Political geographies that are functioning governments range from villages and American Indian reservations with fewer than 100 people to cities, counties, and states with millions of people. Statistical geographies, developed by the Census Bureau with input from users, also range in population size—from blocks with fewer than 10 people to large metropolitan areas. Each functioning government has a need for census data, particularly if the political unit is too small in population to benefit from the American Community Survey (ACS). Figure 11.1 demonstrates the small size of three important political geographies: incorporated places (19,483 in every state but Hawaii); minor civil divisions (MCDs—17,870 in 21 states); and American Indian reservations and trust lands (337 in 40 states). Statistical geographies such as census designated places (12,134 in 50 states, with a median population of about 700 people) and census tracts (83,848 populated tracts in 50 states and the District of Columbia, with a median population of about 3,775 people) are also important for many data users.[25]

The small population size of many geographies makes it challenging to inject noise while providing acceptable accuracy; the task is harder yet because many geographies are not nested in a neat hierarchy. While blocks nest into block groups, census tracts, counties, and states, other types of geographies are "off-spine" (see Figure 11.2).[26] A handful of American Indian reservations cross state boundaries, and almost 40% cross county boundaries. While MCDs nest within counties, 7% of incorporated places cross county boundaries. Moreover, 40% of incorporated places cross census tract boundaries, as do many MCDs (see Census Geographies Project, 2022:32).

[24]Similarly, researchers frequently use public use microdata samples from the census (and surveys) to construct dependent and independent variables to analyze household and family characteristics—for example, differences in home ownership for households of various sizes as related to the race-ethnicity of the household. The separate person and housing privacy-protected microdata files for 2020 (see Table F.1) preclude the ability to carry out such studies unless the household variables of interest to the researcher are already encoded in the records of household heads.

[25]Data from the 2020 P.L. 94-171 Redistricting File; see the state tables at the Census Geographies Project, https://mdi.georgetown.edu/census-geographies-project/.

[26]The use of the term "spine" instead of "nested" can imply that "off-spine" geographies are less worthy of accurate data, which is not the case given that many off-spine geographies are functioning governments and locally recognized communities.

Copyright National Academy of Sciences. All rights reserved.

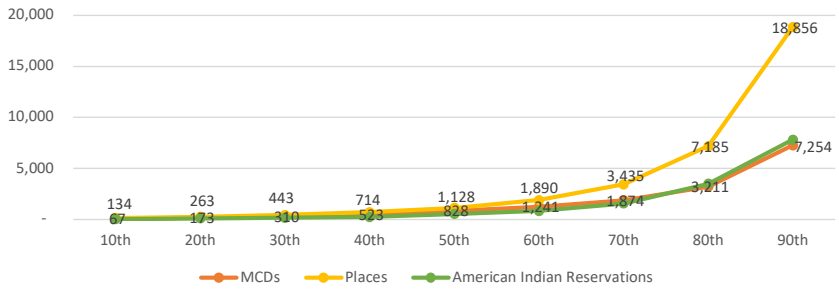


Figure 11.1 Population at the 10th–90th percentiles, minor civil divisions, incorporated places, and American Indian reservations, 2020 Census.

NOTES: MCDs, minor civil divisions. Population numbers not shown for American Indian reservations and off-reservation trust lands because they are too similar to MCDs to be legible.

SOURCE: Census Geographies Project (2022), Figure 3-4, middle panel, using data from 2020 Census Redistricting File.

Unlike the problems of invariants and joins, the Census Bureau did not appear to fully recognize the off-spine problem until December 2019, when users presented comparisons of 2010 SF1 tables as originally released with TDA-protected 2010 SF1 tables at a Committee on National Statistics workshop (National Academies of Sciences, Engineering, and Medicine, 2020). These comparisons showed much larger differences for geographies that did not nest compared with those that did. The Census Bureau then worked to "optimize" the spine and in other ways ensure greater accuracy for off-spine geographies.

11.3.3 Implications for Development of the 2020 Disclosure Avoidance System

By late 2018, as Version 4 of the 2020 operational plan announced that the 2020 Census DAS would use differential privacy, the Census Bureau was implementing its TDA for redistricting data from the June 2018 End-to-End Census Test. The Census Bureau and external critics also began assessing the planned methods on the 1940 Census complete count file, which was available through IPUMS at the University of Minnesota.

In the first half of 2019, calls grew for more formal testing of the new DAS before final implementation (see, e.g., Ruggles et al., 2019). Discussions

Copyright National Academy of Sciences. All rights reserved.

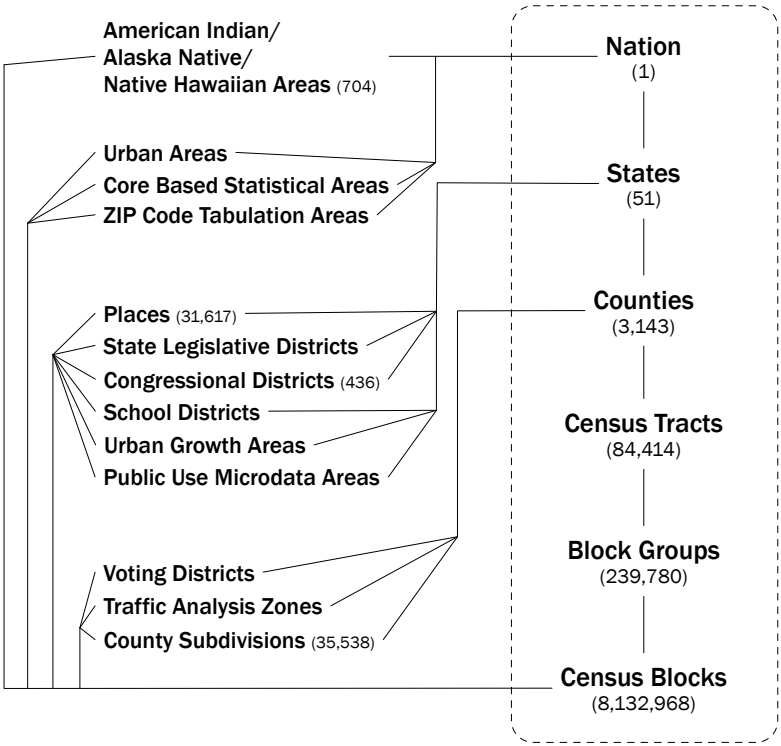


Figure 11.2 Hierarchy of census geographic entities.

NOTES: Entities in the dashed rectangle are the "spine" or "on-spine" geographic levels, in the parlance that has developed around the Census Bureau's TopDown Algorithm; those not in the rectangle are "off-spine." Counts associated with geographic levels are tallies from the 2020 Census Redistricting File, as derived by Census Geographies Project (2022).

SOURCES: Adapted from "Standard Hierarchy of Census Geographic Entities" published by U.S. Census Bureau at https://www2.census.gov/geo/pdfs/reference/geodiagram.pdf?#.

Copyright National Academy of Sciences. All rights reserved.

at a CNSTAT workshop held in June 2019[27] supported the need for more testing with the opportunity for user feedback (see Hotz and Salvo, 2022, for a chronicle of events during this period). Beginning in fall 2019, the Census Bureau responded by releasing "demonstration products" of noise-infused data from the 2010 Census which could be compared with the official published 2010 releases.[28] Each release was followed by an opportunity for user feedback.[29] The Census Bureau also offered webinars and developed materials to inform users about the new DAS (see Section 11.4 for more about the Census Bureau's communication efforts and our assessment of those efforts in relation to user needs). Key decisions resulting from the demonstration file releases and associated user feedback (including feedback from subject matter specialists at the Census Bureau) included:

- To continue and refine the TDA for the Redistricting and DHC Files, which meant continued separate processing of person and housing tables;[30]
- To prioritize development of the DAS for the Redistricting File, given the more pressing statutory deadlines for apportionment and redistricting data (and the lack of motion in being granted an extension on those deliverables), and to slow active DAS development for the DHC File;
- To adopt discrete Gaussian rather than discrete Laplace noise infusion to reduce the prevalence of outliers with large differences between the demonstration products and the originally published statistics;
- To tune the noise infusion in the Redistricting File so that the largest demographic group in any area of more than 500 people would differ no more than ±5 percentage points from the collected value 95 percent of the time (responding to Voting Rights Act requirements on majority-minority districts);[31]
- To optimize the geographic spine, including the creation of a separate spine for American Indian or Alaska Native (AIAN) areas by state, to improve accuracy for the many off-spine geographies;
- To develop separate algorithms (SafeTab-P and SafeTab-H) for the DDHC-A and B files that could produce better-quality data than the TDA

[27]See https://www.nationalacademies.org/event/06-06-2019/challenges-and-new-approaches-for-protecting-privacy-in-federal-statistical-programs-a-workshop.

[28]The Census Bureau also eased user concerns that differential privacy-based algorithms would soon be applied to ACS data products by announcing in summer 2019 that no such change would be made before 2025 at the earliest (Jarmin, 2019b).

[29]See Table F.2 in Appendix F.2 for a list of releases beginning with the 2018 End-to-End Redistricting File and related events for the DAS development from 2018–2023.

[30]The Census Bureau experimented with a "bottom up" approach with the 1940 Census data, in which noise was injected in all of the smallest geographic areas (enumeration districts) in one operation but quickly ruled out this approach because the noise-infused data would become less rather than more accurate when aggregated for larger geographic units. See Abowd (2019a).

[31]See Wright and Irimata (2021) for an empirical analysis.

Copyright National Academy of Sciences. All rights reserved.

Table 11.2 Features of the TopDown Algorithm Compared with the SafeTab-P, SafeTab-H, and PHSafe Algorithms

TopDown Algorithm—2020 Redistricting, DHC, Demographic Profile Files	SafeTab-P, SafeTab-H, PHSafe Algorithms—2020 DDHC-A, B; Supplemental-DHC Files
Produces privacy-protected microdata	Produces privacy-protected tabulations directly
All geographies add up as expected	No requirements for geographies to add up
When aggregating data, noise generally cancels out and data become more accurate	Aggregating data increases noise the more one aggregates
Consistency across data products	Not consistent with Redistricting or DHC Files
Overall accuracy can be targeted but exact accuracy levels cannot be known in advance	All margins of error determined in advance and met 95% of the time
Same amount of data provided for each type of geography	Detail of data provided depends on population/household size of geographic area or race/ethnicity group (Census Bureau terms this "adaptive design" but it is similar to cell suppression used in censuses prior to 1990)

NOTES: DHC, Demographic and Housing Characteristics (File); DDHC, Detailed DHC (Files).

SOURCE: Adapted from Devine et al. (2023:Slide 21).

for the large number of race, ethnicity, and AIAN tribes and villages, although at the cost of inconsistency with the Redistricting and DHC Files (see Table 11.2 for a summary of the differences between the TDA and the new algorithms);

- To develop yet another algorithm (PHSafe) for the Supplemental-DHC File that could handle complex household-person joins; and
- To successively increase the value of the privacy-loss budget (ϵ) for tables in the Redistricting File from 4 (persons) and 2 (housing units) in the 2010 Demonstration File, October 2019 version, to 17 (persons) and 2.5 (housing units) in the final file (released August 2021)—these increases responded to user feedback after each demonstration product release that identified large differences for key variables compared with the original 2010 data.[32] The Census Bureau increased the value of ϵ for

[32]The Census Bureau has never set a global value for ϵ for the suite of 2020 Census data products, although differential privacy theory calls for setting such a budget and then allocating the budget among data products and then among variables and geographies in each product. As noted in Section 11.2, epsilon (ϵ) is used to express the privacy-loss budget although rho (ρ) is the parameter in the discrete Gaussian implementation (see Appendix F.1) because users became accustomed to ϵ values.

Copyright National Academy of Sciences. All rights reserved.

the combination of person and housing unit tables in the Redistricting and DHC Files from 40 (2010 DHC Demonstration File, March 2022 version) to 53 in the final DHC File (released May 2023).

User comments on successive demonstration products commended the Census Bureau for improved accuracy over earlier products but continued to raise concerns about remaining anomalies, particularly for small off-spine geographies and small race and ethnic groups (see Table F.2). The Federal State Cooperative on Population Estimates Steering Committee (2020) asked that "illogical and implausible values" evident in the demonstration products be addressed before final parameters were set. "At the root of these issues," they wrote, "is the separation of people and housing in the DAS processing." They suggested adding invariants, which would have put more constraints on the already-stressed DAS implementation and were not adopted.[33]

In August 2021, when the official 2020 Redistricting File was released, the "illogical" counts remained, and the larger public learned of the impact of the DAS on small-area counts. The Census Bureau recommended aggregating blocks to address the issue but did not otherwise provide detailed guidance to users. At least two private-sector data-dissemination firms are offering the Redistricting File to users with anomalous data smoothed or otherwise altered (Cassal, 2022; Hodges and Cortes, 2022).

11.4 ASSESSMENT

We assess the development and implementation of the 2020 DAS in three ways. First, we refer to statements of principles, practices, and fundamental responsibilities for federal statistical agencies; second, we contrast the DAS development process with other census innovations; and third, we briefly look at how other countries are responding to contemporary threats to confidentiality protection. We also note the effects of the 2020 DAS on other Census Bureau programs, including DA and population estimates, and on the Census Bureau's ability to produce informative substate operational quality metrics (e.g., self-return rates) for public use.

We do not pass judgment on the appropriateness of differential privacy-based algorithms in the toolkit of statistical confidentiality protection. We are aware of effective applications and are confident there will be more such as the theory and available suite of algorithms evolve. The issue before us is the application of differential privacy-based algorithms for 2020 Census products.

[33]For example, the committee recommended: "In addition to adding occupied housing units to the list of invariants please: • Limit the distribution of household population to blocks with occupied housing units [and] • Force each occupied housing unit to have at least one person associated with it" (Federal State Cooperative on Population Estimates Steering Committee, 2020).

Copyright National Academy of Sciences. All rights reserved.

11.4.1 Fundamental Responsibilities of Statistical Agencies

Box 11.5 reproduces relevant text from three well-known and highly regarded statements of statistical agency responsibilities. All three statements put the production of relevant, timely, accurate, and objective statistics for the public good as the first responsibility of statistical agencies. Each list also includes protection of confidentiality of respondents' information but as an instrument to enable the agency to provide high-quality data by gaining respondents' cooperation.

Principles and Practices for a Federal Statistical Agency (National Academies of Sciences, Engineering, and Medicine, 2021:156) also introduces the useful concept of data stewardship in its description of the operational principles and best practices established in the U.S. Office of Management and Budget's Federal Data Strategy.[34] It states that the Federal Data Strategy practices, grouped under the rubrics of "Building a Culture that Values Data and Promotes Public Use," "Governing, Managing, and Protecting Data," and "Promoting Efficient and Appropriate Data Use," "represent aspirational goals that are intended to improve the government's approach to data stewardship and the leveraging of data to create value."

The Census Bureau's original decision to develop the 2020 DAS using formal privacy methods did not contradict the various lists of principles (although it contradicted the tenets of census planning—see Section 11.4.2). Its decision to continue with that approach, however, when it became clear that it would be difficult, if not impossible, to produce timely, accurate, relevant data for users, arguably did. The continuation decision prioritized confidentiality protection above data utility broadly defined.

The situation in which no product beyond the 2020 Redistricting File had been released until just over three years after Census Day is a marked departure in fulfilling reasonable (based on previous censuses) expectations of data users for timely data release. That the scheduled release dates for the DDHC-B File (counts of household types and tenure by detailed race, ethnicity, and AIAN tribal and village groups) and the Supplemental DHC (S-DHC) File (household-person join tables) have been pushed back to September 2024 means that data release will not be complete until almost the end of year four after Census Day.

With regard to accuracy, users have documented problems with the Redistricting and DHC File demonstration products that remain despite the Census Bureau's best efforts to fine tune the TDA and the decision to increase the privacy-loss budget to levels that were previously believed to be disclosive. With regard to relevance, the paucity of person-household join tables relative to 2010 and earlier censuses and their limited availability for geographic areas

[34]See Memorandum M-19-18, Federal Data Strategy—A Framework for Consistency, at https://www.whitehouse.gov/wp-content/uploads/2019/06/M-19-18.pdf.

Copyright National Academy of Sciences. All rights reserved.

Box 11.5 Statements of Statistical Agency Responsibilities

Foundations for Evidence-Based Policymaking Act of 2018, Sec. 3563, Statistical Agencies [derived from Statistical Policy Directive No. 1, U.S. Office of Management and Budget, 2014; codified at 44 U.S.C. § 3563(a)(1)]

Each statistical agency or unit shall—

(A) produce and disseminate relevant and timely statistical information;
(B) conduct credible and accurate statistical activities;
(C) conduct objective statistical activities; and
(D) protect the trust of information providers by ensuring the confidentiality and exclusive statistical use of their responses.

United Nations (2014) *Fundamental Principles for Official Statistics*

Principle 1 Official statistics provide an indispensable element in the information system of a democratic society, serving the Government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honor citizens' entitlement to public information.

Principle 2 To retain trust in official statistics, the statistical agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data.

Principle 3 To facilitate a correct interpretation of the data, the statistical agencies are to present information according to scientific standards on the sources, methods and procedures of the statistics.

Principle 4 The statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics.

Principle 5 Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents.

Principle 6 Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.

Principle 7 The laws, regulations and measures under which the statistical systems operate are to be made public.

Principle 8 Coordination among statistical agencies within countries is essential to achieve consistency and efficiency in the statistical system.

Principle 9 The use by statistical agencies in each country of international concepts, classifications and methods promotes the consistency and efficiency of statistical systems at all official levels.

Principle 10 Bilateral and multilateral cooperation in statistics contributes to the improvement of systems of official statistics in all countries.

Copyright National Academy of Sciences. All rights reserved.

Box 11.5 (continued)

Committee on National Statistics *Principles and Practices for a Federal Statistical Agency* (National Academies of Sciences, Engineering, and Medicine, 2021)

Principle 1: Relevance to Policy Issues and Society Federal statistical agencies must provide objective, accurate, and timely information that is relevant to important public policy issues.

Principle 2: Credibility Among Data Users and Stakeholders Federal statistical agencies must have credibility with those who use their data and information . . . through clear public commitments to professional practice and transparency in all that they do, including informing users of the strengths and weaknesses of their data.

Principle 3: Trust Among the Public and Data Providers Federal statistical agencies must have the trust of those whose information they obtain [by protecting confidentiality and respecting data providers].

Principle 4: Independence from Political and Other Undue External Influence Federal statistical agencies must be independent from political and other undue external influence in developing, producing, and disseminating statistics.

Principle 5: Continual Improvement and Innovation Federal statistical agencies must continually seek to improve and innovate their processes, methods, and statistical products to better measure an ever-changing world.

is problematic for many census data users, who find such tables essential for government services planning and other purposes.

A benefit of formally private approaches is their transparency in many respects. For approaches like data swapping, the percentage of swapped households in an area and other parameters are kept confidential. By comparison, the Census Bureau has released the code and values of key parameters of the 2020 DAS (e.g., the allocation of privacy-loss budget ϵ to tables at various geographic levels). The Census Bureau has also released what are called noisy measurement files (NMFs) for the 2010 demonstration products and the 2020 production Redistricting and DHC data. These files—the result of noise infusion into every table cell—enable analysts to get some sense of the effects of the discrete Gaussian noise infusion per se versus the postprocessing steps in the TDA to eliminate negative values and enforce consistency of the data for smaller geographic areas with larger geographic areas.[35]

[35]Kenny et al. (2023:1) analyzed the 2010 NMF and demonstration files, finding "that the NMF contains too much noise to be directly useful alone, especially for Hispanic and multiracial populations. TopDown's post-processing dramatically reduces the NMF noise and produces similarly accurate data to swapping in terms of bias and noise. These patterns hold across census geographies with varying population sizes and racial diversity. While the estimated errors for both TopDown and swapping are generally no larger than other sources of Census error, they can be relatively substantial for geographies with small total populations." Hopefully, others will replicate

Copyright National Academy of Sciences. All rights reserved.

This technical information about the 2020 DAS is invaluable for advancing the field of confidentiality protection using formally private methods. But there remains a lack of the promised transparency in important senses. The code and technical parameters do not help the vast majority of data users interpret and work with the 2020 data products, nor do they explain the effects of invariants and other deviations from the pure application of differential privacy on the actual confidentiality protection obtained by the 2020 system. There are also lingering questions in the absence of full methodological documentation for a key input to the decision to adopt and stick with the new DAS approach—namely, the 2016–2018 reconstruction and reidentification experiment.

Beginning in late 2019, the Census Bureau came to realize that users needed not only more explanatory information, but also materials that would help them assess the 2010 demonstration files to provide informed feedback and get ready to work with the 2020 data products. Box 11.6 illustrates efforts by the Census Bureau to fill these information gaps, citing examples of webinars, blogs, demonstration products, summary metrics, published papers, handbooks, and issue briefs.

Despite the Census Bureau's efforts, the materials available to date do not address users' needs for (a) concrete guidance and the equivalent of confidence intervals for working with the data; and (b) an understandable metric for estimating the increased disclosure risk with the higher values of ϵ adopted for the production 2020 Redistricting and DHC Files, which would inform users' views on the confidentiality-utility tradeoff. The materials are either too complex for most users to analyze (e.g., the NMFs) or they are too general. For example, the practical advice in the primer released with the 2020 Redistricting File (U.S. Census Bureau, 2021e) boils down to two suggestions: (1) aggregate blocks; and (2) compare the most recent 2010 Redistricting File demonstration product with the originally released 2010 file as a guide to the likely accuracy of the 2020 file. Similarly, users do not know what to make of the allocation of privacy-loss budget ϵ to various topics and geographies except in the most general sense.

The Census Bureau recently announced an initiative to expand its guidance on using the 2020 Census differential privacy-protected data products, to provide measures of total error introduced by the 2020 DAS and mechanisms for generating confidence intervals, and to develop other aids for users (Devine et al., 2023:Slides 16–18). As noted in Box 11.6, several short documents, written at a general level, were released March 27, 2023, describing why the Census Bureau chose differential privacy, the TDA, and disclosure avoidance for the Redistricting File.[36] For estimating error from the privacy-protected

and expand on this analysis to produce a better understanding of the effects of formally private confidentiality-protection techniques on census data.

[36]See https://www.census.gov/programs-surveys/decennial-census/decade/decennial-publications/2020/census-briefs.html under "2023."

Copyright National Academy of Sciences. All rights reserved.

Box 11.6 Types and Examples of Informational Materials Developed by the U.S. Census Bureau for the 2020 Disclosure Avoidance System

- *2010 Demonstration Files*—Redistricting and Demographic and Housing Characteristics Files produced from 2010 data using the TopDown Algorithm (TDA) for disclosure protection. Users can compare the demonstration files with the actual 2010 data products to assess the effects of the TDA on accuracy. First demonstration file released October 2019 (see Appendix Table F.2 for a complete listing of demonstration files and related products released to date).
- *2010 Demonstration File Summary Metrics*—Excel spreadsheets containing such error measures as mean absolute percentage error for tabulations (e.g., total population and race and ethnicity populations for specific types of geographic areas) comparing a demonstration file with the underlying 2010 data. The summary metrics are largely measures of central tendency with a few measures of outliers.
- *Crosswalks*—Excel spreadsheets comparing tables provided in 2020 data products with tables provided in 2010 data products. The first version was published in fall 2019; the latest version is dated May 21, 2023. The various versions are not at one url.
- *Blogs*—Census Bureau leadership and technical staff regularly announced developments for confidentiality protection for 2020 data products in blog posts (see Note below).
- *Webinars*—Census Bureau staff regularly held webinars on confidentiality protection for 2020 data products (see Note below).
- *Primer/Handbook* (U.S. Census Bureau, 2021e)—published November 2021, with the Population Reference Bureau.
- *Frequently Asked Questions*—Maintained at https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-updates/2020-das-faqs.html.
- *Issue Briefs* (with the Population Reference Bureau)—Three briefs released on March 27, 2023: Disclosure Avoidance and the 2020 Census Redistricting Data: What Data Users Need to Know (https://www2.census.gov/library/publications/decennial/2020/census-briefs/c2020br-02.pdf; Disclosure Avoidance and the 2020 Census: How The TopDown Algorithm Works (https://www2.census.gov/library/publications/decennial/2020/census-briefs/c2020br-04.pdf); Why the Census Bureau Chose Differential Privacy (https://www2.census.gov/library/publications/decennial/2020/census-briefs/c2020br-03.pdf).
- *Papers*—Census Bureau staff published several papers on the 2020 DAS in refereed journals, including Abowd and Hawes (2023); Keller and Abowd (2023); Abowd et al. (2022); Abowd and Schmutte (2019).
- *Presentations*—Census Bureau staff regularly made presentations at meetings of the Census Scientific Advisory Committee and National Advisory Committee and at professional conferences and other venues on the 2020 DAS.

NOTE: Many of the above items and other materials can be found at the Census Bureau's "Understanding Differential Privacy" page at https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/differential-privacy.html. See also https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/protections-by-data-product.html.

Copyright National Academy of Sciences. All rights reserved.

2020 data products, the Census Bureau is investigating an "approximate Monte Carlo" or "parametric bootstrap" approach. This approach is computer-resource-intensive but, if the Census Bureau can work through the technical and computing issues, it could well provide useful metrics to assist users in understanding the error in 2020 data products.[37]

Vink (2023), New York's state representative to the Federal-State Cooperative for Population Estimates, transformed summary metrics provided by the Census Bureau for the Redistricting and DHC Demonstration Files into fitness-for-use scores (e.g., mean absolute percentage error or MAPE < 10% = 1, MAPE > 10% = 0). The individual scores can be combined into an overall fitness-for-use score for a variable. He illustrated these scores using age in the August 2022 and October 2019 DHC 2010 Demonstration Files, concluding that the August 2022 file was substantially improved over the original October 2019 file, but that age distributions for small-population geographic units were still problematic. His approach may be useful for other variables while waiting for the Census Bureau to provide tools for estimating error bounds.

11.4.2 Census Planning, Testing, and Innovation

The decennial census is the U.S. government's largest peacetime operation and has tight deadlines, mandated in law, for the delivery of reapportionment and redistricting data, in addition to expectations for timely release of additional data generated from enumerating a large and highly diverse population. The planning process (see discussion of the 2020 operational plan series in Section 11.3) is lengthy, detailed, and meticulous, and involves extensive testing and research, which typically spans multiple censuses. Sometimes, the process has taken longer than need be,[38] but, generally, robust experimentation and testing, including under the conditions of an actual census, are very much in order for major census innovations. Examples include the Mailout/Mailback census introduced for two-thirds of housing units in the 1970 Census, after several decades of smaller-scale testing (National Research Council, 2010) and the multidecade development of the modern Post-Enumeration Survey (PES) methodology for coverage measurement (National Research Council, 2009). In contrast, the decision to adopt a markedly different and much more complex method for confidentiality protection was made at virtually the eleventh hour of census planning, without prior testing in the census context or buy-in from the user community.

Problems in the rollout of new methods and technologies are not unexpected. Virtually every census has encountered the unexpected from external

[37]See U.S. Census Bureau (2023c:Slides 71–102); see also Neunhoeffer et al. (2023).

[38]For example, the Census Bureau ran out of time to develop the logistics for a second questionnaire mailing in 2000 by conducting its own tests on response-rate effects despite decades of positive survey research findings (National Research Council, 2010).

Copyright National Academy of Sciences. All rights reserved.

forces (e.g., COVID-19 in 2020 and the failure of a key contract in 2010). Time and again, the Census Bureau has shown outstanding ability to modify affected operations in real time to cope with such exigencies. The decisions to adopt the new DAS very late in planning for 2020 and not to reconsider in the face of serious implementation challenges, however, were the Census Bureau's own doing. Moreover, the Census Bureau did not have a backup plan were the new DAS to prove more difficult and time consuming to implement than expected.

The Census Bureau consults often with users and stakeholders about census planning through advisory committees, congressional hearings, presentations at professional conferences, interactions with National Academies' study panels, U.S. Government Accountability Office staff, and others. For decisions that affect data products in major ways, however, such engagement is rarely enough. Extensive and repeated user interactions are essential to identify key user needs and obtain buy-in when painful trade-offs are needed between, in this case, confidentiality protection versus data availability and utility.

Although many Census Bureau staff interact frequently and effectively with data users, the organization apparently did not realize the incompatibilities between differential privacy-based methods, as developed prior to their adoption for the 2020 Census, and user needs for official (invariant) counts, household-person join variables, and accuracy for many small, off-spine political geographies. The Census Bureau made every effort to accommodate those needs, but its solutions are not yet viewed as satisfactory by users, and the costs in delays and undermining of user trust have been great.[39]

The comment periods following the release of each demonstration product update (see Table F.2) gave users the opportunity to provide feedback that the Census Bureau used to improve the next update. Many users with limited resources, however, found it difficult to meet the short, frequent deadlines, and many others were not aware of the commenting opportunities (O'Hara, 2022).

In addition to using differential privacy-based methods, the Census Bureau cut back on the content of many data products, either by eliminating tables or limiting them to larger geographic areas than in previous years. Such actions are a reasonable way of enhancing confidentiality protection, when used proportionately in concert with other masking or noise infusing methods and when done in consultation with users. The Census Bureau, however, made most of these decisions on its own. For example, the very first "crosswalk" the Census Bureau issued in fall 2019 proposed decreasing the geographic detail in the DHC File for a number of tables compared with their counterparts in 2010. The final DHC File restored this detail for many of the affected tables.[40] The Census

[39]See Garner (2023). Abowd and Hawes (2023:23–24) acknowledge underestimating the challenges of effective user communication.

[40]This information had been maintained in a "Change Log" tab in the Data Table Guide at https://www2.census.gov/programs-surveys/decennial/2020/program-management/data-table-guide-dhc-dp.xlsx, though it seems to have been dropped from recent versions.

Copyright National Academy of Sciences. All rights reserved.

Bureau's recent decision to drastically reduce geographic detail (eliminating data for all substate geographies) for the already-trimmed set of household-person join tables in Supplemental-DHC File (U.S. Census Bureau, 2023b) was not discussed in advance with data users.

The Census Bureau also could not avail itself of a potential opportunity to cut back on the race detail provided for redistricting, which could have significantly increased privacy protection for that product without impairing utility. Its late start ruled out that opportunity, given that the redistricting community of federal and state agencies, state legislatures and redistricting commissions, the political parties, and the consulting firms that draw up redistricting plans would need to be fully onboard with any potential changes.

11.4.3 Privacy Protection in Other National Statistical Offices

The United Kingdom (UK)'s Office of National Statistics recently investigated use of differential privacy-based methods on mortality data and concluded (Dove, 2021:19):

> The independent noise addition method is best suited to releases with a limited set of outputs, known ahead of time. ... The top-down method we attempted to apply suffered from significant bias issues, arising from perturbing small counts as well as perturbing zeros. Perturbing zeros increases the noise given and causes additional information loss (less utility). Assigning proportionally more epsilon to lower levels in the hierarchy also slightly reduced utility. This was possibly because as the adjustment to higher level totals were performed sequentially, higher level totals are more important. Assigning more epsilon to high levels may slightly improve results.

The UK adopted a varied suite of confidentiality-protection methods for its 2021 census,[41] including swapping, cell suppression, and the cell key method developed by the Australian Bureau of Statistics.[42] The UK has a threshold of at least 100 people for the smallest geographic areas ("output areas") for which it releases census data, which would seem restrictive. As it turns out, the UK has never seen the need for smaller geographic areas—it does not have many

[41]See "Protecting personal data in Census 2021 results" at https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/methodologies/protectingpersonaldataincensus2021results.

[42]See Templ and Sariyar (2022:1236). The cell key method "adds noise to frequency and contingency tables using predefined look-up tables consisting of random number values, which allows producing consistent results from queries on dynamically generated tables. Before the tables are generated, a fix[ed] numerical code is assigned to every record. All codes of records falling in a cell are summed, and this sum, the cell key, is used for selecting the random number from the look-up table." The protection is done on the source microdata so that data products can roll out without further protection needed.

Copyright National Academy of Sciences. All rights reserved.

small-population governmental jurisdictions nor does it have legal and court requirements for redistricting that require block-level data.[43]

Bach (2022) reviewed random noise approaches for statistical disclosure control for official censuses and other population statistics conducted by members of the European Union. He observed (Bach, 2022:669):

> - In most countries, censuses are very costly investments that must be justified by their unique value added for policymakers and researchers. Consequently, user concerns must be taken very seriously. . . .
> - For typical systematic attacks on static outputs . . . strictly ϵ-DP mechanisms (with unbounded noise) do not offer significant benefits over utility-driven approaches with bounded noise. . .
> - Mechanisms based on unbounded noise . . . may ruin unique utility features of censuses, such as small-area accuracy.

Bach (2022:678) noted that statistical disclosure control is the responsibility of member states—Eurostat will not receive or house confidential data. For protection of the 2021 round of censuses, the European Statistical System (see Antal et al., 2017) therefore recommended the use of bounded noise methods, particularly the cell key method, for protection of outputs with limits on the amount of noise injected. Bach (2022:671) stressed the importance that official statistical agencies "invest more in explaining clearly and exhaustively what certain noise protection setups imply for . . . products."

11.4.4 Effects on Other Census Bureau Programs

The adoption of the 2020 DAS led to delays and changes in methodology for key Census Bureau programs, including:

- *Demographic Analysis:* To date, DA estimates of net coverage error for the 2020 Census are available for single years of age and sex but not for Black and all other races. The latter are essential to extend the time series of differential net undercount estimates between race groups, which are a key high-level indicator of trends in census quality. The reason is a longer-than-anticipated development of a privacy-protected "Modified Race File," which would permit direct comparisons of census counts and

[43]For example, England has 309 local authorities (compared with over 40,000 functioning local governments in the United States), with a median population of about 143,000. The three smallest have 2,000, 8,600, and 41,000 people, respectively; the largest has 1.1 million people. See "Demography and migration data" at https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/demographyandmigrationdatacontent/2022-11-02#demography-unrounded-population-estimates and the description of geographies in the UK 2021 Census at https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeographies/census2021geographies.

Copyright National Academy of Sciences. All rights reserved.

DA estimates for Black people and all other races (see Chapter 10 for details).

- *Population Estimates for the Nation, States, and Counties:* The population estimates program, mandated in Title 13, also requires a Modified Race File (see Chapter 10). Given the delays in obtaining such a file and to keep to its annual schedule for updates, the population estimates program adopted a "blended base." The estimates currently start with 2020 Census counts for the nation, state, and counties protected with a small amount of noise infusion; incorporate estimates of race/ethnicity by age and sex starting with the *2010-based* estimates updated through 2020 and forward; and control the estimates at the national level using 2020 DA age and sex estimates (see Hartley, 2022).

Another example of a Census Bureau program for which decisions appear to be driven, at least in part, by the nature of the differential privacy-based mechanism for the 2020 Census data is defining urban areas based on housing unit criteria rather than population criteria.[44]

11.4.5 Effects on Operational Quality Metrics

The Census Bureau deserves credit for releasing operational quality metrics (e.g., percentages of occupied households that self-responded or for which administrative records were used, and percentages of people missing responses for individual items) soon after completion of the census. This was not the historical practice but was prompted by user requests, given the uncertainty about census quality due to the COVID-19 pandemic and other factors that could have affected response.[45]

Operational metrics provide indirect indicators of differences in quality for areas for which reliable estimates of coverage errors are not available from the PES or DA. For example, self-responses are known to be more accurate in terms of coverage and content than responses obtained by proxy in Nonresponse Followup (NRFU) (see Chapters 3, 4). This knowledge can help users interpret the data and can help the Census Bureau work with communities to improve the next census.

[44]See "Redefining Urban Areas following the 2020 Census," https://www.census.gov/newsroom/blogs/random-samplings/2022/12/redefining-urban-areas-following-2020-census.html. Housing unit counts are invariant, but occupancy status and population are not.

[45]See https://www2.census.gov/programs-surveys/decennial/2020/data/operational-quality-metrics/operational-quality-metrics-technical-documentation.pdf. Since 2000, the Census Bureau has released unperturbed, real-time response rates for states, counties, and census tracts as a spur to respond. These metrics, however, are limited for quality assessment because the denominator is all addresses in the self-response universe (including addresses that turn out to be vacant or nonresidential/nonexistent).

Copyright National Academy of Sciences. All rights reserved.

The Census Bureau issued four releases of 2020 Census Operational Quality Metrics in 2021–2022, providing comparable metrics (to the extent possible) for the 2010 Census.[46] Release 1 in April 2021, Release 2 in May 2021, and Table 2 of Release 3 in August 2021 covered data-collection modes for addresses, occupied versus vacant housing units by census operations, and item nonresponse rates; these data were protected by rounding or noise infusion. However, states are too internally heterogeneous for operational metrics to be of much utility in assessing quality. Table 1 of Release 3, also in August 2021, provided summary statistics (means, medians, standard deviations) for counties and census tracts within states for selected metrics (e.g., percent addresses responding by internet or proxy), and these data were also protected by rounding. Finally, in October 2022, the Census Bureau's Release 4 included 6–8 operational metrics for individual counties and tracts, using differential privacy with $\epsilon = 3$ for disclosure avoidance. Unfortunately, these metrics, particularly for tracts, are too noisy for use, as can be seen by adding up the percentages of NRFU enumerations by a household member, a proxy, or administrative records within a tract. Representing the three principal response modes in NRFU, the three percentages should roughly add to 100% but vary widely around that mark.[47]

On metrics, our interim report (National Academies of Sciences, Engineering, and Medicine, 2022) stated:

> **[Interim Report] Conclusion 4.6:** It will not be possible for this panel (or any other evaluator) to understand and characterize the quality of the 2020 Census unless the Census Bureau is forthcoming with informative data quality metrics, including new measures based on operational/process paradata, at substate levels and small-domain spatiotemporal resolution, unperturbed by noise infusion. . . .
>
> **[Interim Report] Recommendation 4.1:** The Census Bureau should work on ways to make 2020 Census data quality metrics publicly available at small-domain spatiotemporal resolutions, unperturbed by disclosure avoidance, to bolster confidence in the published tabulations. The Census Bureau should also develop ways to enable qualified researchers to access a full range of data quality metrics and report their findings.

[46] All of the Operational Quality Metric releases in public form are available under "Evaluating Quality" at https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/data-quality.html.

[47] C. Dick, Demographic Analytics Advisors (2023, January), Operational Metrics—A Focus on Release 4 and NRFU. PowerPoint presentation to the Census Quality Reinforcement Task Force. The Census Bureau also has not yet included small-area "low response scores" or other operational metrics from the 2020 Census—including Self-Response return rates (self-responses from occupied housing units)—in its Planning Database, presumably because it is working out the privacy-protection issues. This file provides invaluable information for local groups striving to encourage response to the census and other purposes. See "Planning Database" at https://www.census.gov/topics/research/guidance/planning-databases.html.

Copyright National Academy of Sciences. All rights reserved.

The Census Bureau justified a low ϵ for operational quality metrics because of a JASON recommendation that the Census Bureau not spend any of its privacy-loss budget on substate metrics.[48] Yet the Census Bureau has never provided a convincing argument as to how disclosure of such metrics for small areas increases identification risk for individuals in the census.[49] While theoretically there could be a risk, it is difficult to imagine a plausible attack scenario. Importantly, not publishing such metrics closes off a way for data users to determine the confidence to place in data for specific areas and hampers research into improvements for 2030.[50]

11.5 CONCLUSIONS

Conclusion 11.1: In an era of Big Data, linkage technology, and other aspects of today's computational and data environment, it is difficult for the decennial census to balance the need for accurate data for small areas and small population groups with an adequate level of confidentiality protection. When evaluating methods and levels of confidentiality protection, stewardship of census data for the public good requires balancing societal needs for the information—particularly for equity in representation, fund allocation, and many other uses—with plausible harms of breaches of confidentiality.

Conclusion 11.2: The decision of the U.S. Census Bureau to respond to the threats to confidentiality protection at a very late date in 2020 Census planning with a new, more complex Disclosure Avoidance System (DAS) using differential privacy-based algorithms went counter to long-standing principles of decennial census planning. The approach had not been tested in a census environment nor had the ability of the algorithms to handle critical user data needs been assessed; the Census Bureau had no backup plan should implementation of the new DAS prove challenging. The decision to continue to deploy the new DAS in the face of serious implementation problems has resulted in marked delays in delivery of data products, with some variables and

[48]JASON (2021:8). JASON is a group of scientists who conduct analyses for the U.S. Department of Defense and other agencies; their activities are managed through the MITRE Corporation.

[49]The 2010 Census provided item imputation rates for blocks in SF1 and census tracts in SF2—see Table F.1 in Appendix F.2.

[50]The CSAC recommended, May 25, 2021, "that the Census Bureau not apply DP [differential privacy] to substate quality metrics that are being released to the public. … Applying DP to the quality metrics could make them largely irrelevant, and will take part of the PLB [privacy-loss budget] away from important future data products. If the Bureau concludes that the quality metrics cannot be released without applying DP, CSAC requests that the Bureau justify their decision by explaining how these metrics could be used in a reconstruction scenario." See https://www2.census.gov/about/partners/cac/sac/meetings/2021-05/2021-05-25-census-response.pdf.

Copyright National Academy of Sciences. All rights reserved.

types of geographic units of questionable utility, and other variables and geographies not provided at all. In addition, it is not clear that the chosen privacy budgets for the various 2020 Census data products, with high values of the ϵ parameter that trade off accuracy with confidentiality protection, provide much actual protection.

Conclusion 11.3: Sustained, interactive communication with data users in advance of major changes to data products from a census or long-running survey is critical for a statistical agency to maintain credibility with and serve its user community. The late decision to adopt new confidentiality-protection techniques for the 2020 Census led to untimely and inadequate user communication on the part of the U.S. Census Bureau. The comment periods after the release of each demonstration file for the Redistricting and Demographic and Housing Characteristics Files were valuable, but many users lacked the staff and resources to respond adequately.

Conclusion 11.4: By not releasing unperturbed operational quality metrics for census tracts, the U.S. Census Bureau missed an opportunity to inform the public of quality differences in the 2020 Census among substate areas and to gain the assistance of interested stakeholders in planning for 2030. Other than the theoretical concept that the release of any piece of information about an individual increases the risk of disclosure, the Census Bureau has not justified its decision to release census tract metrics that are too noisy for meaningful analysis.

Conclusion 11.5: For a major change in data products such as the new confidentiality-protection techniques for the 2020 Census, it is essential for a statistical agency to document fully and publicly the need for the change and the implications for the products. In the case of the 2020 Census, there has not yet been peer review of the crucial 2016–2018 reconstruction-reidentification study that was the public basis for the decision to adopt a formal privacy approach for protecting confidentiality at a late date in 2020 Census planning. Similarly, there has not been an actual demonstration of why more-than-minimal confidentiality protection is needed for substate operational quality metrics (e.g., self-return rates for census tracts and counties) or why it is not possible to provide as-collected population counts to local governments with a reasonable level of confidentiality protection for other data.

Copyright National Academy of Sciences. All rights reserved.

11.6 LOOKING AHEAD TO 2030

The use of differential privacy-based algorithms for the 2020 Census data products is a *fait accompli*. That said, it would be imprudent for the Census Bureau to assume that such algorithms will be appropriate or sufficiently well developed for 2030 data products, as seems to be the assumption in the Census Bureau's fiscal year 2024 budget submission.[51] The differential privacy framework is predicated on protecting against any attack no matter how unlikely the scenario, injecting noise into every cell and every item in datasets save for those quantities that are held invariant.[52] "Tuning" can increase data utility for some use cases at the cost of decreasing utility for others. These features make the approach useful for one-off applications, such as protecting the output from a researcher's use of confidential data in a Federal Statistical Research Data Center, but challenging when it comes to providing widespread, equitable data access to multipurpose, multidimensional data sets such as the decennial census or the ACS. Information providers, such as Google and Microsoft, have used differential privacy-based algorithms for confidentiality protection, but their data products do not need to inform redistricting, fund allocation, or program planning for small geographies and population groups and do not need to deal with an essentially unbounded set of queries.

In December 2022, the Census Bureau issued the following statement about confidentiality protection for the ACS, acknowledging that the differential privacy approach is a work in progress for censuses and surveys intended to serve a wide public (Daily, 2022):

> Our current assessment is that the science does not yet exist to comprehensively implement a formally private solution for the ACS. We expect a multiyear development period, including data user review and feedback, that will extend beyond 2025. . . . It's also not clear that differential privacy would ultimately be the best option. Other formally private disclosure avoidance approaches may end up being a better fit for the ACS.

11.6.1 Framework for Balancing Utility and Confidentiality

Given the reality of contemporary threats to confidentiality protection alongside the imperative for statistical agencies to serve the public, the best way forward remains in question. The Foundations for Evidence-Based Policymaking Act of 2018 and the Year 2 Report of the Advisory Committee on Data for Evidence Building (ACDEB, mandated by the Foundations Act)

[51]See https://www.commerce.gov/sites/default/files/2023-03/Census-FY2024-Congressional-Budget-Submission.pdf, p. 157.

[52]This text was revised after the prepublication release in order to clarify previous wording on the applicability of noise infusion, regardless of whether particular variables might popularly be deemed noninvasive.

Copyright National Academy of Sciences. All rights reserved.

Box 11.7 Relevant Language—Foundations for Evidence-Based Policymaking Act of 2018

The Foundations for Evidence-Based Policymaking Act of 2018 added to the Confidential Information Protection and Statistical Efficiency Act to require statistical agencies to, "to the extent practicable," "expand access to data assets [to] develop evidence while protecting such assets from inappropriate access and use," tasking the U.S. Office of Management and Budget with issuing regulations to enable statistical agencies to carry out those requirements. The regulations (not yet issued) are to include:

- Standards for categorizing sensitivity levels of statistical agency data assets and the corresponding levels of accessibility;
- Minimum and maximum numbers of sensitivity levels;
- Requirements to determine the feasibility of "removing or obscuring information in such a manner that the identity of the data subject is *less likely to be reasonably inferred* [emphasis added] by either direct or indirect means;"
- Requirements to conduct comprehensive risk assessments of data assets; and
- Requirements for agency assessments to be "transparent and easy to understand."

SOURCE: Codified at 44 U.S.C. § 3582.

include useful language and frameworks for making nuanced decisions about disclosure risk and data-utility tradeoffs (see Boxes 11.7 and 11.8). The language in the Foundations Act refers to risk assessments and reasonableness of reidentification, and the ACDEB Year 2 Report has a list of guiding principles that bring a framework of reasonableness to the confidentiality-utility debate. Hotz et al. (2022) call for cost-benefit analysis in decisions about an appropriate DAS for a census or survey, with explicit consideration of the loss to society from data that are unusable or only marginally useful due to noise injection.

The Education Sciences Reform Act of 2002 (see Box 11.9) provides language that speaks to shared responsibility of statistical agencies and users to protect privacy by penalizing *any person* for disclosing confidential information, which would include staff of other government agencies, the private sector, and indeed any user, as well as statistical agency staff. The National Center for Education Statistics cites this language on its website for specified datasets. It could be possible to expand the applicability of the Education Sciences Reform Act to all federal statistical data as an amendment, for example, to the Foundations Act.

Laws of member states that implement the European Union's General Data Protection Regulation of 2018 also provide relevant language. For example, Section 171 of the UK Data Protection Act of 2018, "Re-identification of de-identified personal data," states "(1) It is an offence for a person knowingly or recklessly to re-identify information that is de-identified personal data

Copyright National Academy of Sciences. All rights reserved.

Box 11.8 Excerpts—Year 2 Report, Advisory Committee on Data for Evidence Building

Recommendation 1.6. **OMB [The U.S. Office of Management and Budget] should adopt a *risk-utility framework* [emphasis added] as the basis for standards on sensitivity levels, access tiers, and risk evaluations as part of the regulation on expanding secure access to CIPSEA [Confidential Information Protection and Statistical Efficiency Act] data assets.**
Risk and utility should be separately measured but jointly determined. Evidence on data use should be used to inform the measurement of value, models from the public and private sectors should be used to measure risk, and the Five Safes framework should be applied to develop combined risk-utility metrics that are open and transparent.

Recommendation 1.8. **OMB, in coordination with the ICSP [Interagency Council on Statistical Policy], should identify models for shared responsibility among data providers and users and provide guidance on applying such models through the regulation on expanding secure access to CIPSEA data assets.**
The Committee affirms the importance of establishing standards to assess data assets in terms of sensitivity levels, corresponding levels of accessibility, and whether less sensitive versions of the data can be created, as required by [44 U.S.C. § 3582]. The Committee concludes that the regulation should reflect a set of guiding principles, including that *(1) disclosure risk is on a continuum and is not binary, (2) not all data are equally sensitive, (3) there is shared responsibility between the statistical agency and users for protecting and not disclosing or re-identifying data, and (4) there is a need to protect good faith actors (i.e., data providers and users who take all precautions appropriate for known risks).* [emphasis added]

SOURCE: Advisory Committee on Data for Evidence Building (2022:28,34).

Box 11.9 Legislation on Shared Statistical Agency–Data User Responsibility for Privacy Protection

The Education Sciences Reform Act of 2002 directs the Director of the Institute of Education Sciences, U.S. Department of Education, to "develop and enforce standards designed to protect the confidentiality of persons in the collection, reporting, and publication of data under this title." The act holds that "*no person* may—A. use any individually identifiable information furnished under this title for any purpose other than a *research, statistics, or evaluation* purpose [or] B. make any publication whereby the data furnished by any particular person under this title can be identified." [emphases added] Critically, the section on violations holds that:

> *Any person* who uses any data provided by the Director, in conjunction with any other information or technique, to identify any individual student, teacher, administrator, or other individual and *who knowingly discloses, publishes, or uses such data for a purpose other than a statistical purpose,* [or otherwise violates these provisions] shall be found guilty of a class E felony and imprisoned for not more than five years, or fined [or] both. [emphases added]

SOURCE: Codified at 20 U.S.C. § 9573, and reprinted in the Institute of Education Sciences *Restricted-Use Data Procedures Manual*, Appendix D.

Copyright National Academy of Sciences. All rights reserved.

without the consent of the controller [e.g., statistical agency] responsible for de-identifying the personal data."[53]

We do not underestimate the complexity of risk-utility analysis or claim the ability of legislation to deter any and all attacks on the confidentiality of federal statistics. The alternative, however, of attempting to use a single technical approach to protect an agency's data products, chosen for its purported ability to guard against every kind of attack, no matter how implausible, is unlikely to enable a statistical agency to fulfill its mission to provide useful information.

11.6.2 Data Product Plan for 2030

Even as 2020 data products are being developed, it is not too early to begin consultations with users on their data needs for the 2030 Census and to stand up research projects to plan for a timely rollout of 2030 data products. A post hoc analysis could help in developing the 2030 data-products strategy. The Federal State Cooperative on Population Estimates Steering Committee (2022) urged the Census Bureau:

> To commission an independent assessment to evaluate: the Census Bureau's decision processes leading up to the adoption of DP methods for the 2020 Census; its interactions (or lack thereof) with users; the soundness of its reconstruction/ reidentification studies, which formed the basis for the move to DP methods and reduction of content; and evidence on the timeliness, accuracy, and relevance (or lack thereof) of the privatized data.

Census Bureau director Robert Santos' reply to the FSCPE letter said the Census Bureau was "seriously considering this recommendation."

A useful addition to the evaluation outlined above would be an in-depth assessment of which tables and geographies omitted from the 2020 Census products users would argue to reinstate, and which turned out to be inessential. A discussion with the redistricting community about the race/ethnicity detail on the 2030 Redistricting File is critical, in light of the proposal by the U.S. Office of Management and Budget to adopt a combined race/ethnicity question with Middle Eastern and North African and Hispanic categories (see Chapter 10). A reduction in race/ethnicity detail on the 2030 Redistricting File would pay dividends in confidentiality protection.

Two areas of research for the Census Bureau to pursue vigorously leading up to 2030 are candidate disclosure-protection methods, including differential privacy-based algorithms, and risk-utility, cost-benefit analysis for determining confidentiality-utility tradeoffs. For this work, the Census Bureau could consider issuing one or more challenges to enlist the academic and private-sector research communities, in addition to contracting for specific projects. The Census Bureau could partner on challenges with the National Institute of

[53] See https://www.legislation.gov.uk/ukpga/2018/12/section/171/enacted.

Copyright National Academy of Sciences. All rights reserved.

Standards and Technology (NIST).[54] An important element of such research would be constructing understandable measures of risk and accuracy for confidentiality-protected data for each candidate method.

Another area of research would be in-depth assessment of the sensitivity of specific types of information in the census to a range of communities. Such research could consider, for example, whether household relationship and type were appropriate to tabulate at the block level as in 2010, given the differing views in the country about same-sex relationships, or whether to provide that information at a higher level of geography. Similarly, assessment of the actual disclosure risk of reporting unprotected operational metrics at the census tract level could be conducted.

Finally, bolstering communications expertise and resources could have considerable payoff for a post-mortem analysis of 2020, and on research looking toward 2030 in terms of user needs, data sensitivity, cost-benefit analysis, and confidentiality-protection methods. Adequate resources will be essential beginning in the next fiscal year for the Census Bureau to undertake this multiprong program. If the program is well supported, it could improve not only the 2030 Census, but also other Census Bureau programs and the entire federal statistical system to the benefit of policymakers and the public.

Recommendation 11.1: For 2030 Census data products, the U.S. Census Bureau should adopt the risk-utility framework recommended by the Advisory Commission on Data for Evidence Building, which accepts that disclosure risk is a continuum, that not all data items are equally sensitive, and that federal statistical data need to be accessible and useful for a wide range of users and uses.

Recommendation 11.2: At a minimum, the 2030 Census should provide as-collected (i.e., unaltered) total population counts for all governmental units (states, counties and equivalents, minor civil divisions, incorporated places, recognized tribal areas, school districts) and quasi-governmental units (census county divisions, census-designated places, and tribal statistical areas), no matter how small. In addition, census block population totals should add up to block groups, census tracts, counties, and states.

Recommendation 11.3: For the 2030 Census data product plan, the U.S. Census Bureau should begin immediately on a multipronged research program with ample testing and opportunities for feedback and dialogue with the data user and

[54]See Ridgeway et al. (2021) and the overview of NIST differential privacy challenges at https://lish.harvard.edu/nist-differential-privacy-challenges.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

stakeholder community, broadly defined. The goal should be an end-to-end plan by 2027–2028 for producing a suite of 2030 data products that serve user needs, are appropriately protected, and meet the time schedule of the 1990–2010 Censuses.

Recommendation 11.4: U.S. Census Bureau research and dialogue with the user community, broadly defined, on the end-to-end plan for the 2030 Census suite of data products should include:

- **A review of the 2020 Census experience, including rigorous assessment of the Census Bureau's reconstruction-reidentification studies under plausible scenarios of potential attack and comprehensive debriefing of users.**
- **Consultation with the redistricting community on race/ethnicity and geographic detail (accounting for the possibility of a combined race/ethnicity question) that best trade off accuracy and confidentiality protection for the 2030 Redistricting File.**
- **Issuance of challenges to the research community, in addition to focused contracts, for research and development on a range of confidentiality-protection methods and understandable metrics of risk and accuracy to accompany those methods.**
- **Research on practical methods for users to account for noise injected into 2030 Census data by the selected confidentiality-protection techniques.**
- **Research on the application of cost-benefit and risk-utility analysis for making tradeoffs between confidentiality protection and utility of census data products.**
- **Research on the sensitivity of individual data items to relevant communities at differing levels of geographic aggregation and the implications for confidentiality protection.**

Recommendation 11.5: The U.S. Census Bureau should welcome initiatives to add language to appropriate legislative vehicles, such as the Foundations for Evidence-Based Policymaking Act of 2018, that prescribes responsibilities and penalties for data users, in addition to agency staff, for willful, harmful disclosure of confidential information.

Copyright National Academy of Sciences. All rights reserved.

– 12 –

Learning from 2020, Preparing for 2030

The U.S. Census Bureau overcame massive challenges and withstood immense pressures in carrying out the 2020 Census. The delays and operational changes induced by the COVID-19 pandemic strained the process of census-taking—and were arguably even more injurious to the Post-Enumeration Survey (PES), a fundamental tool for ascertaining census quality—but did not break the process. The 2020 Census was not perfect, to be sure; group quarters (GQs) are challenging to enumerate in any census and were even more so in 2020, and the diminished quality of proxy responses (as evident in the striking age-heaping phenomenon) calls the very utility of proxy reporting into question. But it is unrealistic to expect perfection in any census. The Census Bureau should rightly take pride in having a solid plan for the 2020 Census, adhering to that plan as much as possible, and making careful and reasonable adaptations as necessitated by the pandemic.

Yet it serves no one well to downplay a very important point: some challenges were forced upon the 2020 Census but others were 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) with an entirely new approach, one that had not been tested, prototyped, or deployed in the population census context. While confidentiality protection is a critically important responsibility of a statistical agency, this decision was made without appropriate consideration and balance regarding the utility of resulting census data products to fulfill the many important functions of census data. In short, the new DAS was not ready

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

for use in 2020 Census production, and substantially degraded the value of 2020 Census data products in terms of both quality and availability. Moreover, lingering questions about both the simulated database reconstruction attack that motivated the new DAS and the degree of confidentiality protection that was ultimately realized in practice have arguably harmed 2020 Census data products and the Census Bureau's reputation.

Going forward, to best build a research and evidentiary base for conducting the 2030 Census, we echo and strongly endorse the core guidance of our predecessor Panel to Review the 2010 Census (National Research Council, 2011): to effect major change in census conduct, the Census Bureau should focus its primary attention on a small number of carefully selected major innovation areas and conduct a rigorous program of strategic testing and systems development to address the challenges of *how* to implement those innovations rather than *whether* to do so. The predecessor panel articulated the four major innovation areas—field reengineering, response options (including internet response), administrative records, and geographic resource improvement—that were adopted by the Census Bureau in framing the 2020 Census. We believe that the capacities and efficiencies associated with these four priority areas were crucial to the successful conduct of the 2020 Census under supremely adverse conditions, as they each helped the Census Bureau to adapt to the difficult conditions of 2020 as best could be expected.

With this chapter, we close this report by presenting our detailed recommendations from the preceding chapters in a unified framework, along with some additional and overarching suggestions. These recommendations generally fall under three general headings: a suggested set of priority goals for 2030 Census research and development, along with other overall recommendations for improving census evaluation; suggested changes for 2030 Census procedures and operations that are still very important to overall census quality even if they fall outside the top-priority tier; and improvements in the legal and regulatory environment in which the decennial census operates.

12.1 PRIORITY GOALS FOR 2030 CENSUS RESEARCH AND DEVELOPMENT

The predecessor National Research Council (2011:Recommendation 3) Panel to Review the 2010 Census suggested that the Census Bureau adopt a broad motivational goal: "to reduce significantly (and not just contain) the inflation-adjusted per housing unit cost relative to [the 2010 Census], while limiting the extent of gross and net census errors to levels consistent with both user needs and cost targets." To progress toward that goal, that panel suggested that the Census Bureau focus its research and testing efforts on the four innovation areas previously mentioned in Section 2.1. We concur that

Copyright National Academy of Sciences. All rights reserved.

this approach should continue into 2030 Census planning, and the list of key objectives that we suggest has important points of overlap and divergence with last decade's guidance. Specifically, we recommend as follows (note that there is no rank ordering of importance in this list):

Recommendation 12.1: To prioritize its testing and development of both procedures and technical systems for the 2030 Census, the U.S. Census Bureau should take the following as primary objectives:

- **Maximize self-response to the census, including better matching of contact and communication strategies to the desired response mode, with particular attention to hard-to-reach, at-risk populations;**
- **Improve the quality of data in Nonresponse Followup, including reducing if not eliminating the use of low-quality proxy reporting when there is an alternative available;**
- **Reduce gaps in coverage and data quality associated with race, ethnicity, and socioeconomic status;**
- **Improve the quality of address listings and contact strategies for all living quarters, including group quarters; and**
- **Realign the balance between utility, timeliness, and confidentiality protection in 2030 Census data products.**

12.1.1 Maximize Census Self-Response

The "maximize self-response to the census" plank of Recommendation 12.1 is, admittedly, an inevitable reaction to the experience with promoting census self-response—particularly via the internet—in other countries. The U.S. 2020 Census achieved an Internet Self-Response rate of about 62–63% of occupied housing units,[1] which was a formidable achievement, but one need only look at the progression to 80%-plus online response in some recent international censuses (see Table 12.1) to wonder what is possible in 2030. But, by this goal, we mean something more ambitious than it may seem at first blush, and more than just hitting a higher internet-uptake rate in the 2030 Census. Rather, the challenge going forward is to cast the traditional type of enumeration area (TEA)/contact-strategy assignment as less a mechanical exercise in questionnaire delivery and more as a call to action, tailoring that call as much as possible to specific target audiences for maximum effect. There is some portion of the population that will readily reply to the census, and the goal is to make those people aware of the Internet Self-Response option as quickly

[1]Panel calculation from (U.S. Census Bureau, 2021a).

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Table 12.1 Response Rate (%) Via Online Forms in Recent International Censuses

Census Year	Australia	New Zealand*	Canada	United Kingdom–England and Wales
2011	34.3	34.0	53.9	16.4
2016	58.8	87.9	68.3	
2021	78.9		84.1	88.9

* New Zealand's 2011 Census was canceled and postponed to 2013 by the Christchurch earthquake, shifting the usual 5-year cycle by 2 years—2016 to 2018 and 2021 to 2023. Data collection in the 2023 New Zealand Census was delayed by cyclone activity.

SOURCE: Harding et al. (2022:Table 3.3.2).

and efficiently as possible. There is another part of the population that will ever be reluctant, if not completely refuse, to participate, and the challenge for that segment is to try to find the messaging and modes of delivery that could achieve the best results. There are many gradations between these extremes. The Census Bureau's work in delineating TEAs (Self-Response, Update Leave, and Update Enumerate) in 2020, and particularly in subdividing Self-Response into Internet First and Internet Choice, was solid. A challenge for 2030 is revisiting the whole nexus of TEA delineation—from information delivery/contact strategy, to the sequence of mailings and information broadcasts, to basic communication strategy and messaging—to efficiently steer potential respondents to the mode of choice (or the mode most likely to result in completion). This might involve more nuanced approaches than a simple Internet First/Internet Choice divide, a further shift from Update Leave to Self-Response, and—for the truly difficult to count—an in-person initial contact akin to Early Nonresponse Followup or Update Enumerate, if that is most likely to yield results.

Specific recommendations consistent with this theme include:

> **Recommendation 6.1: The U.S. Census Bureau should continue to investigate the performance of the contact strategies used in the 2020 Census, to increase overall self-response (particularly via the internet) in 2030, more in concordance with international experience with online census response. This investigation should include review of the effectiveness of the 2020 Census's segmentation into Internet First and Internet Choice strategies, as well as the timing and the operative contact/enumeration strategy (e.g., Internet First or Internet Choice mailings or Update Leave delivery) underlying successful Non-ID enumerations in 2020.**

Copyright National Academy of Sciences. All rights reserved.

Recommendation 6.2: The U.S. Census Bureau should engage in further research to more specifically identify social, economic, and housing characteristics associated with less-than-accurate forms of enumeration, and should also research related communication and operational strategies that would improve self-response and associated data accuracy for those population and housing segments associated with poorer-quality census data in 2020.

Recommendation 6.3: The U.S. Census Bureau should assess the magnitude of the impact of imputation on the accuracy of 2020 Census data. Assessment of the differential impact of imputation on subpopulations should be used to prioritize research into strategies to improve self-response.

Recommendation 5.5: The U.S. Census Bureau should research and test approaches to increase the proportion of dwellings that can be mailed to while still being associated with a physical geographical location. This would reduce the proportion of addresses in Update Leave and expand the use of approaches yielding higher Self-Response for more dwellings in the 2030 Census.

12.1.2 Improve Quality of Nonresponse Followup Data

Census quality is a very complex construct, and it can be subject to dangerous oversimplification. We adopt the basic premise that self-response to the census is preferred over other modalities because it is likely to provide higher-quality information but there are no absolute guarantees. A census return is not necessarily perfectly accurate by dint of being returned by mail or transmitted with a keystroke; answers can still be missing or misleading. Likewise, administrative records data can vary in quality according to their source and recency, and neighbors or landlords contacted to provide proxy responses might have great familiarity with the households in question—or such responses might not be materially better than guesswork. Thus, even with attempts to maximize Self-Response, Nonresponse Followup (NRFU) will remain an important part of 2030 Census data collection, and so it is natural to ask how NRFU data might be collected more accurately—in terms of both coverage accuracy and content completeness/accuracy.

There are many possible dimensions for improvement of NRFU data that the Census Bureau can and should explore, including reassessing the minimum and maximum number of visits to a nonresponding housing unit or the number of unsuccessful visits before resorting to alternative means; or considering whether centralized telephone follow-up could improve efficiency

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

when accurate phone contact information is available. Our analyses of data on NRFU resolutions and the marked degree of age heaping in the 2020 Census consistently suggest that proxy reporting in 2020 NRFU was particularly problematic. For reasons that are not yet fully understood, proxy returns were particularly diminished in quality in 2020 compared with 2010, and this makes the assessment of proxies particularly important. Much could be learned in census testing about the nature of proxy response and trying to differentiate between more reliable and less reliable proxy responders—to try to attach a measure of confidence in the information provided during a proxy interview. Proxy-related questions should also be considered for research as part of a more detailed interrogation of 2020 Census operational data. But, in terms of broader visions for the 2030 Census and beyond, it is worth considering whether proxy reporting ought to be reserved for extraordinary circumstances—if permitted at all—if high-quality alternative information is available.

Specific recommendations along these lines include:

Recommendation 7.1: As a goal for the 2030 Census, the U.S. Census Bureau should consider a major reduction in the use of proxy interviewing for enumeration, if not the elimination of proxy reporting in all but very limited circumstances. Work toward this goal should be predicated on the results of research that only the Census Bureau can conduct, reexamining and comparing proxy reports in the 2020 Census with the information that would have been available from administrative records or third-party data resources.

Recommendation 7.2: The U.S. Census Bureau should consider conducting research on the quality, and factors associated with variation in the quality, of proxy and count-only responses, with the aim of minimizing the occurrence of these factors.

Recommendation 7.3: The U.S. Census Bureau should conduct fine-grained analysis of Nonresponse Followup enumerations (separately for household, proxy, and administrative records responses) for small geographic areas and population groups as part of its research to develop and test ways to improve the quality of enumeration in the 2030 Census. The Census Bureau should publish the results of its analysis as an aid to users of 2020 Census data.

12.1.3 Reduce Quality and Coverage Gaps by Race, Ethnicity, and Socioeconomic Status

In considering various approaches to enumeration and ways to optimize their implementation, it is important for the Census Bureau to pay special

Copyright National Academy of Sciences. All rights reserved.

attention to outcomes for traditionally hard-to-count populations, such as racial and ethnic minorities, those with lower socioeconomic status, and babies and young children. Our analyses show clear patterns of less-optimal response modalities (i.e., lower Self-Response and higher NRFU proxies) for these groups, and the PES shows increases in differential undercoverage relative to 2010 for racial and ethnic minorities and for renters.

The challenge for 2030 Census planning is not just identifying these kinds of quality and coverage deficits through analysis of 2020 Census and other data, but developing specific, implementable actions to address the deficits. One clear direction in this regard is for the Census Bureau, in conjunction with the U.S. Office of Management and Budget and with Congressional oversight, to reach closure early on the mechanics of collecting race and ethnicity data in 2030. The reporting of race and Hispanic origin in the 2020 Census was partially impaired by the relatively late-stage continuation of the two-question format for these items on the questionnaire itself, coupled with changes in the coding and processing of write-in categories. Early closure on the form of the relevant race and Hispanic-origin questions to be used in 2030 is certainly not a panacea for quality and coverage differentials, but is a good step nonetheless.

Along these lines, we recommend:

Recommendation 10.1: The U.S. Census Bureau should conduct research to determine how the changes in format and processes that were made in the 2020 Census and in the American Community Survey beginning in 2020 affected the distributions of race and ethnicity. Such research should use qualitative, quantitative, and simulation methods to ascertain: how respondents viewed and used the 2010 and 2020 formats; trends in multirace reporting by age, sex, race, and ethnicity; how samples of 2020 respondents would have been categorized using the 2010 format, data capture, and coding; and the implications of differences in write-ins by response mode (e.g., more write-ins for internet responses) on race distributions among geographic areas. The Census Bureau should communicate the results of this research to data users to assist them to understand the implications of the changes for time series, and should also use the findings to inform 2030 Census planning.

Recommendation 10.2: To improve the quality of reporting on race and ethnicity in the census, American Community Survey (ACS), and other federal data collections, the U.S. Office of Management and Budget (OMB) should revise Statistical Policy Directive No. 15 to adopt a combined, check-more-than-one race/ethnicity classification with both Hispanic and Middle Eastern or North African as main categories, in addition to

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

White, Black, American Indian and Alaska Native, Asian, and Native Hawaiian and Other Pacific Islander. OMB should formalize any changes to the race and ethnicity reporting standards as soon as practicable, to permit the U.S. Census Bureau to implement the new race/ethnicity question in the ACS and in the 2030 Census. OMB should allow agencies to collect data only for the major categories in the combined question when added detail (as in the 2020 Census) would impose undue administrative burdens. A decision by the Census Bureau on including added detail and write-in spaces, along with expanded data capture and coding procedures, for a combined race/ethnicity question for the ACS and the 2030 Census should await completion of research on such matters as how respondents viewed the 2020 question format and the effects of response-mode differences in write-in responses on diversity among geographic areas.

Recommendation 10.3: The U.S. Census Bureau should produce a crosswalk or bridge between the 2010 and 2020 Census race and ethnicity questions and responses. Similarly, if a combined race/ethnicity question (with Hispanic and Middle Eastern or North African categories) is adopted as the standard by the U.S. Office of Management and Budget, the Census Bureau should produce a crosswalk or bridge between the 2020 version and revised race and ethnicity questions and responses, as soon as the revisions are implemented in the American Community Survey and then in the 2030 Census. The Census Bureau should involve data users in this important work.

Recommendation 10.4: The U.S. Census Bureau should consult as early as possible with the redistricting community, which includes state legislators, redistricting commissions, political parties, and political consultants, among others, to determine the optimum set of tabulations to include on the 2030 Redistricting File, whether the current race and ethnicity categories are retained or the proposed revisions take effect. The consultation should include consideration of streamlining the number of race/ethnicity categories in the file and combining blocks with small populations (using input from localities) to maximize the ability to protect confidentiality while maintaining accuracy.

Copyright National Academy of Sciences. All rights reserved.

12.1.4 Continuous Modernization of Geographic Resources for All Living Quarters

An understated success story of the 2020 Census is that there was not a comprehensive, listers-walking-every-street Address Canvassing operation in 2019. Rather, through partnership work early in the decade and development of systems for in-office review of imagery and other data resources, the Census Bureau shifted the great bulk of Address Canvassing work to in-office review rather than in-field listing. This fourth plank of Recommendation 12.1 echoes the "modernizing geographic resources" innovation area suggested by National Research Council (2011) and urges the Census Bureau to continue its good work in the geographic arena, but also calls for escalation in important ways. First, In-Office Address Canvassing (IOAC) should endure—early, often, and continuously—but it should also be allowed to fulfill its original vision. The early discontinuation of Active Block Resolution due to budget constraints was regrettable, as it diminished IOAC's role to one of simply documenting discrepancies in housing unit counts between the Master Address File (MAF) and available imagery (rather than actually correcting address issues through in-office research). The seemingly solid performance of the Office-Based Address Validation operation (verifying respondent-supplied addresses in census returns without a MAFID through query of in-office data) portends good things about a fully realized IOAC program. Second, the small but very relevant nonhousehold population living in GQs warrants commensurate attention earlier in the decade; the Census Bureau has taken first steps toward making the MAF an ongoing repository of all living quarters, household and GQ alike, and these efforts should continue. Finally, in both the household and GQ context, an address repository is necessary but not sufficient; attention should be paid to the best ways to actually contact residents at those addresses (e.g., by mail or in person) and collect that contact-strategy information on more than a once-a-decade-refresh basis.

Accordingly, we recommend:

> **Recommendation 5.4: Lessons learned from the Office-Based Address Verification operation should be documented and incorporated into the tools and techniques for ongoing, continuous address and map data refinement (e.g., continuous In-Office Address Canvassing) as well as revisions of the Non-ID Processing operation.**
>
> **Recommendation 9.1: The U.S. Census Bureau's steps to make the Master Address File an ongoing, continuous inventory of all living quarters—including group quarters (GQs)—should continue, with an eye toward developing updated information for GQs on par with the regular updating of conventional**

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

housing units. This should include a separate inventory of GQ types that have longstanding address provenance and those that need to be tracked more closely because they are continually changing.

Recommendation 9.2: Group quarters (GQs) require more than updated address information; they require continually updated contact strategies as well. The U.S. Census Bureau should develop "selected advance contact" protocols with GQ facilities by type (and their parent companies, as appropriate), akin to follow-ups performed with businesses, and state, local, and tribal partners should be engaged in this update activity.

Recommendation 9.3: In light of the 2020 experience, the U.S. Census Bureau should convene group quarters (GQ) stakeholders in a discussion of more effective and useful options for electronic provision of data for GQ residents, and on the issues involved in obtaining access to administrative records for GQ facilities. The completeness of such records should also be assessed by GQ type.

Recommendation 5.3: The U.S. Census Bureau should extend and improve during-the-decade programs, akin to the Geographic Support Systems Initiative of the 2010s, with the goal of continuous address- and map-database refinement. Such efforts should include engagement with state, local, and tribal partners, including on the question of whether some addresses can safely be removed from the Master Address File.

12.1.5 Realign Balance between Utility, Timeliness, and Confidentiality Protection in 2030 Census Data Products

As discussed in Section 1.2.2, census quality is a multidimensional construct, and unduly weighting any one dimension to dominate all others can be detrimental. One important dimension of data quality—timeliness of data—fell particularly to the wayside in the 2020 Census experience, in part due to delays arising from the COVID-19 pandemic but mainly due to the late-stage implementation of disclosure avoidance based on differential privacy (see Chapter 11), which had the effect of weighting confidentiality protection above all. This fifth plank of Recommendation 12.1 calls for the balance to be righted among critical aspects of quality in 2030.

Both national and international sources identify timeliness as an aspect of data quality. *Principles and Practices for a Federal Statistical Agency* specifies as part of practice 9 that "dissemination should be timely, and information should be made readily available on an equal basis to all users" (National

Copyright National Academy of Sciences. All rights reserved.

Academies of Sciences, Engineering, and Medicine, 2021:93). In both domestic and international standards for statistical quality, timeliness is one of many dimensions that comprise overall quality (Federal Committee on Statistical Methodology, 2020 and Box 1.1).[2]

For most uses of the census, data are time sensitive. Census law in Title 13 of the U.S. Code requires that apportionment counts be delivered by December 31 of Census Year and the P.L. 94-171 redistricting data three months later (by April 1 of Census Year + 1). The time sensitivity of these data is due to the constitutional necessity to reapportion the U.S. House of Representatives and the desire by most states to have recent data for redistricting. Yet there are additional purposes—some governmental and others in the private sector—for which census data are not just time sensitive but actually perishable. Stale data can become useless data.

In prior U.S. censuses, timeliness and geographic disaggregation were both major foci of the decennial census data products program. Through 1990, basic counts for places were reported in "preliminary" counts, by the field staff, very shortly after the field enumeration. This was done as early as May or June of the census year, so that local officials could see and challenge counts (as appropriate) before they were made official, and this was done even for small places.[3] In the 20th century, cross-tabulations were released in "advanced" unbound formats, which were then assembled into hardback books by states, usually by Census Year + 2. (Computer summary files, as described in Chapter 11, were also released on a timely basis.) Moreover, in providing these timely glimpses, the Census Bureau was well aware of the needs of localities. In the introductory description of the "publication program" from the 1960 Census Procedural History, there is emphasis on "timeliness of the data" and the "quantity of 1960 statistics published:" (U.S. Census Bureau, 1966)[4]

---

[2]See also https://unstats.un.org/unsd/methodology/dataquality/references/1902216-UNNQAFManual-WEB.pdf.

[3]See, for example, the account in the *New York Times* describing the August 1990 release of preliminary 1990 Census estimates for about 39,000 local governments in 26 states ("Earliest Figures for Census Fall Below Estimate," August 23, 1990, at https://www.nytimes.com/1990/08/23/us/earliest-figures-for-census-fall-below-estimate.html.

[4]The 1960 Census procedural history continues (U.S. Census Bureau, 1966):

> The quantity of statistics published for the 1960 Censuses was greater than that of earlier censuses. One of the major policy decisions resulting from consultation with the Census Advisory Committees and with the Conferences of Census Users was that more data should be published for small areas, both for use by local communities and for detailed research on variations of characteristics within larger areas. Users of the census data could then group the data for small areas to build statistics for whatever larger area was appropriate for their purposes, such as a homogeneous economic area or the area served by a particular facility. In addition there was an increase in the subject matter for which statistics were published.

Copyright National Academy of Sciences. All rights reserved.

> The goal of the publication program for the 1960 Censuses, like that of previous censuses, was to make the data available as soon as possible after the census was taken, and to do so while maintaining the Bureau's standards relating to the quality of the statistics issues and the way presented.

Other English-speaking countries that conducted traditional censuses in the COVID-19-affected year of 2021—Australia, Canada, and England and Wales—have already released their findings. By comparison, the U.S. 2020 Census has slipped further behind in releasing data—some of which (i.e., missing the December 21, 2020, deadline for apportionment totals) is eminently explainable by the lengthy fieldwork delay induced by the pandemic. But the COVID-19 pandemic is not an adequate explanation for all of the subsequent delays; rather, it appears that the adoption of differential privacy-based disclosure avoidance presented both anticipated and unanticipated obstacles to releasing the data. The Census Bureau itself most forcefully expressed this in its May 31, 2023, press release announcing the delay of some detailed 2020 Census data products until September 2024: "Our existing disclosure avoidance system could not ensure confidentiality protections for people within substate geographies for these tables while also meeting the Census Bureau's data quality standards" (U.S. Census Bureau, 2023b). Moreover, the same press release announced a startling reduction in the number of geographies that would be available in the Supplemental Demographic and Housing Characteristics Files product—from original plans to publish some tables at the "county, census tract, block group, American Indian/Alaska Native/Native Hawaiian [area], and [place]" level to the national and state levels *only*.

Withholding local data is also a departure from previous U.S. census practice. The planned unavailability of most 2020 data for local geographies affects school districts, emergency planners, land-use planners, and many other users. The identification of local needs, and the creation of policies and programs to address those needs, rely on decisions that need to be based on data that accurately reflect the relevant population. The absence of timely population data for small geographic areas, for example, can seriously compromise the ability of local officials to respond to public health crises or, more generally, to prioritize use of scarce resources to support local communities. The delay and reduced detail constitute a markedly diminished return on the nation's investment. Importantly, this is a data-equity issue for nonmetropolitan areas, many of which are already counted as poverty-persistent areas.[5] Small communities are especially disadvantaged, and, in many states, the majority of communities are relatively small. Larger metropolitan areas, big cities, and other large jurisdictions may employ their own demographers and access other local resources for decision making.

[5]See https://www.census.gov/library/stories/2023/05/persistent-poverty-areas-with-long-term-highpoverty.html.

Copyright National Academy of Sciences. All rights reserved.

Moreover, their large size and resources make American Community Survey (ACS) data a viable and timelier alternative. Small towns, villages, and many tribal areas have much more limited choices, given their size, with no real alternatives to decennial census data. This is especially the case for states in which decennial census data are mandated for the distribution of revenue and other resources, leaving very small and frequently vulnerable communities at risk.

We recognize and agree with the importance of confidentiality protection as a major responsibility of the Census Bureau, as a statistical agency. Yet we are concerned that this one dimension of census and data quality has become so dominant in the Census Bureau's approach as to threaten to preclude other critical considerations, such as utility and timeliness. Accordingly, in early planning for 2030, we strongly recommend that order and balance be restored—including a plea for that most basic of census information, the total population count unperturbed by disclosure avoidance, to be made available to functional geographic units.

Recommendation 11.1: For 2030 Census data products, the U.S. Census Bureau should adopt the risk-utility framework recommended by the Advisory Commission on Data for Evidence Building, which accepts that disclosure risk is a continuum, that not all data items are equally sensitive, and that federal statistical data need to be accessible and useful for a wide range of users and uses.

Recommendation 11.2: At a minimum, the 2030 Census should provide as-collected (i.e., unaltered) total population counts for all governmental units (states, counties and equivalents, minor civil divisions, incorporated places, recognized tribal areas, school districts) and quasi-governmental units (census county divisions, census-designated places, and tribal statistical areas), no matter how small. In addition, census block population totals should add up to block groups, census tracts, counties, and states.

Recommendation 11.3: For the 2030 Census data product plan, the U.S. Census Bureau should begin immediately on a multipronged research program with ample testing and opportunities for feedback and dialogue with the data user and stakeholder community, broadly defined. The goal should be an end-to-end plan by 2027–2028 for producing a suite of 2030 data products that serve user needs, are appropriately protected, and meet the time schedule of the 1990–2010 Censuses.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Recommendation 11.4: U.S. Census Bureau research and dialogue with the user community, broadly defined, on the end-to-end plan for the 2030 Census suite of data products should include:

- **A review of the 2020 Census experience, including rigorous assessment of the Census Bureau's reconstruction-reidentification studies under plausible scenarios of potential attack and comprehensive debriefing of users.**
- **Consultation with the redistricting community on race/ethnicity and geographic detail (accounting for the possibility of a combined race/ethnicity question) that best trade off accuracy and confidentiality protection for the 2030 Redistricting File.**
- **Issuance of challenges to the research community, in addition to focused contracts, for research and development on a range of confidentiality-protection methods and understandable metrics of risk and accuracy to accompany those methods.**
- **Research on practical methods for users to account for noise injected into 2030 Census data by the selected confidentiality-protection techniques.**
- **Research on the application of cost-benefit and risk-utility analysis for making tradeoffs between confidentiality protection and utility of census data products.**
- **Research on the sensitivity of individual data items to relevant communities at differing levels of geographic aggregation and the implications for confidentiality protection.**

12.2 IMPROVING CENSUS RESEARCH AND DEVELOPMENT

We will return to specific areas for 2030 Census research-and-development work in the next section, but believe it important to offer some general precepts for such work first.

12.2.1 Continue to Learn from the 2020 Census Experience

The first of these general precepts is simply that there are still many important lessons to be drawn from the 2020 Census—from its as-yet-unreleased data products, from its internal evaluations and assessments, from the vast array of information that electronic data collection made uniquely possible in 2020, and from the history and evolution of 2020 Census procedures and practices.

Copyright National Academy of Sciences. All rights reserved.

Although everyone hopes that the 2030 Census will not be impacted by a global force as disruptive as the COVID-19 pandemic, we are concerned that the 2020 Census could be treated too much as a one-off set of never-to-be-repeated conditions and pressures.

As a first step, we urge the Census Bureau to proceed with its evaluation, experiment, and assessment reports as expeditiously as possible, particularly to the extent that those reports do not simply stop at rote cataloging and instead fuel deeper dives and simulations. In the same vein, a formal, documented procedural history has been invaluable for understanding the development of past decennial censuses. It is true that many pieces can be gleaned from dozens of detailed memoranda, but there is great value in a having a coherent record. The lack of a procedural history for the 2010 Census is unfortunate, but one should be created for 2020.

> **Recommendation 2.1: In addition to completing its own program of evaluation and assessment reports, the U.S. Census Bureau should complete and publish a comprehensive procedural history of the 2020 Census.**

By this precept of not closing the book on the 2020 Census too early, we also acknowledge the time and resource limitations on the work of our panel's data analysis subgroup—and recognize that many of the essential analyses of 2020 Census and PES data can *only* be done by the Census Bureau itself. The Census Bureau should study issues that our panel was unable to examine, and should do so with analyses at subnational spatial and demographic resolution whenever possible. Among the many topics that fall under the heading are: assessment of the reengineered field-operation systems and their impact/efficiency; effectiveness of Non-ID Processing and timing of Non-ID responses; efficacy of differing numbers of NRFU enumeration attempts; and exploration of the impact of individual components of the 2020 Census contact strategy, including whether the effectiveness of the impromptu mailings performed as a result of the COVID-19 pandemic can be assessed.

We encourage the Census Bureau to approach reanalysis of 2020 Census operational and output data, as well as data from the 2020 PES, as analysis of a natural experiment on a grand scale. In many respects, 2020 Census data and paradata offer an ideal opportunity to learn about census procedures, as they were applied under the full trappings of a decennial census and in ways that cannot be replicated well in intercensal tests. In addition, the census itself provides huge sample sizes not available in census tests.

Our analyses in Chapters 6 and 7 illustrate instructive associations between response modes and many ACS variables, at the tract level. Census Bureau analysts have the knowledge and resources to dig much deeper, by incorporating operational variables and performing multivariate analyses (via regression or higher-dimensional tables) to more plausibly approximate causal relationships.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Also, by examining similarities between outlying domains identified through such modeling, the Census Bureau might discover regions or types of housing units not well served by current procedures. Further, the complexity of the census in terms of the variety of data-collection procedures used in different types of geographies for different types of housing units motivates the use of more complicated statistical models to better represent the interaction of different census processes and different types of housing units and types of households.[6]

With specific reference to the PES, we recommend:

> **Recommendation 4.2: The U.S. Census Bureau should conduct additional analyses of the 2020 Post-Enumeration Survey (PES) to learn as much as possible about census omissions. One such analysis should include P-sample and E-sample households with at least one member in common, to assess not only erroneous census enumerations but also omissions and how they vary with household characteristics and type of enumeration (e.g., proxy interview, internet response). In addition, the Census Bureau should perform discriminant or similar analyses to identify variables that contributed to census component errors (e.g., omission, duplication) for 2020 P-sample and E-sample addresses, households, and individuals, which could identify census processes to target for improvement in 2030.**

As a final example, there is likely some useful information contained in the operational paradata that the Census Bureau possibly has retained, concerning difficulties that households experienced during the enumeration process. For instance, what parts of the census questionnaire did individuals have difficulty filling out, based on latencies in response to the internet instrument? What households contacted the Census Bureau via telephone for questionnaire assistance and what languages did those households primarily speak? We are of the impression that paradata remain a useful, underutilized source of information through which the parts of the census processes that were problematic for various subsets of the census universe could be explored.

12.2.2 Improving Census Experimentation and Testing

Over the decades, it has become the norm to think of a census experiment as something that must be done at large scale and that must pit some completely novel approach or technique against standard procedures. Similarly, it has become the norm to equate "census test" with "test census," favoring large-scale dry runs of substantial parts of census data-collection machinery. Both of these

[6]Some candidates for use as independent and dependent variables in such models are provided in Appendix D.

Copyright National Academy of Sciences. All rights reserved.

concepts are limiting in important ways, and we encourage the Census Bureau to approach census experimentation more broadly in planning the 2030 Census and the program of tests that will precede it.

The Census Bureau has regularly conducted randomized experiments as part of intercensal tests of census procedures, and randomized experiments are the most effective way to identify actionable, causal relationships. Accordingly, in the panel's opinion, the Census Bureau should look for opportunities to conduct such experiments to learn about operations under consideration. But we argue that the Census Bureau should also seek opportunities to use randomized experimentation as a learning/evaluation tool. One drawback of 2020 data is that, to our knowledge, there were no substantial randomized experiments that would allow causal inference for any aspects of the enumeration procedures employed in 2020. One example of an opportunity for randomized experimentation was the delineation of Internet First and Internet Choice contact strategies in the 2020 Census. There are still some things to be learned about internet versus paper take-up in the two strategy groups, but insights would be enriched if a small (not disruptive to overall census conduct) experiment with randomized assignment had been embedded in 2020 Census production. Understandably, the Census Bureau's overriding objective in fielding the census is to conduct the best possible (current) count, but we suggest that small provisions for randomizing application of methods could garner much valuable information at very little cost or disruption (and also without severely impacting individuals' ability to be counted in the census). We encourage the Census Bureau to think ahead about embedding such possibilities in 2030 Census production to inform the 2040 Census.

Embedded randomized experimentation might also be the basis for studying the effectiveness and return on investment of the decennial census partnership and communication programs, as suggested in our interim report (National Academies of Sciences, Engineering, and Medicine, 2022:Conclusion 4.5). In the panel's view, planning for the 2030 Census should consider designed evaluations of issues in the formation, coverage, continuity, and timing of these partnership and outreach functions. For a true experiment, one member of a pair of tracts or other domains evaluated as "similar" on relevant attributes could be randomized to intensive outreach, another to less intensive (including no) outreach. Mode-specific participation can be compared, delivering actionable, causal relations. Alternatively, the focus of the experiment could be on ways to increase self-response for groups for which alternatives to the Self-Response mode are weak (e.g, low-quality administrative records sources).

Copyright National Academy of Sciences. All rights reserved.

12.2.3 Systems Design to Facilitate Evaluation

A series of National Academies of Sciences, Engineering, and Medicine census study panels, beginning with National Research Council (1985) and perhaps expressed most forcefully by National Research Council (2004b:§ 8–B), made calls for the decennial census to retain a Master Trace Sample—effectively, a complete audit trail through all census enumeration and follow-up operations for a sample of households that could then be analyzed for cost and error structures. Attempting to shore up the entire technical/information infrastructure of the census, the latter National Research Council (2004b:Recommendation 8.7) panel argued that the vision should be a Master Trace *System* rather than a sample—meaning that this type of audit trail should be available for all households in the census, as a byproduct of regular system functioning.

As we understand it, the Census Bureau made important strides toward this vision in its development of information technology (IT) systems for the 2020 Census and its allowance for operational dashboards by which census managers could monitor real-time flows. Underlying the 2020 Census IT plans was the concept of a Census Data Lake, representing the initial essential step of data and paradata retention. However, the Census Bureau's decision to convert to another software platform—and the subsequent reprogramming effort (see Box 6.2)—made for a late scramble to finalize the systems for deployment.

Accordingly, we echo our predecessor study panels in urging the Census Bureau to develop operational and technical systems for the 2030 Census with the requirement that census systems be amenable to both real-time monitoring and post hoc evaluation and reanalysis. Such a requirement would benefit the understanding of census quality and could ultimately improve all census operations. This recommendation is, in part, a reflection of the panel's difficulty in identifying areas for analysis: our ability to ask for information and paradata was limited because we did not know exactly what information what was being retained and how. We were commonly referred to the Census Bureau's evaluation teams' study plans, which we were grateful to be able to review behind the Census Bureau IT firewall, but those documents had similar questions regarding the data that were actually available because they were generally developed 2–3 years prior. In some instances, our data requests were modified precisely because of the anticipated difficulty in retrieving information from the Census Data Lake without impeding the Census Bureau's ongoing work on 2020 Census production. Technical systems developed with an eye toward facilitating eventual research and evaluation would be a great boon to any evaluator of census methods—most notably to the Census Bureau itself.

Copyright National Academy of Sciences. All rights reserved.

12.3 IMPROVING CENSUS PROCESSES

For a variety of reasons, some topics on which we offer specific recommendations in this report did not rise to the level of priority goals outlined in Recommendation 12.1. This is not to diminish the importance of these topics—they remain vitally important to a quality 2030 Census—and they merit further explication under the general heading of improving census processes.

12.3.1 Continued Work on Administrative Records to Supplement Census Operations

The reason administrative records work is listed here but omitted as a priority goal area in Recommendation 12.1 is expressed in detail in Section 8.6. The way administrative records were used as a supplement to NRFU and other census operations in 2020 was appropriate and careful, including the use of records for enumeration in NRFU instances in which at least one in-person attempt failed to yield a return *and* the Census Bureau had high confidence in the availability of administrative records information for the household in question. There is considerable room for refinement and improvement in this limited, supplementary use of administrative records data, and we encourage it:

> **Recommendation 8.1: The U.S. Census Bureau should continue its research and development program concerning the best ways to use administrative records as a supplement to decennial census operations. Potential uses of administrative records include expanding enumeration of limited subsets of the 2030 Census population in the Nonresponse Followup workload, reducing proxy responses and whole-person imputations, and possibly redressing the long-standing net undercount of children ages 0–4.**

But, as Section 8.6 argues, we do not believe that achieving administrative records enumeration of a very large share of the population, absent any contact attempt, should be a priority goal of 2030 Census planning and implementation. The move to an administrative records-based census is a much longer term proposition for which the evidentiary, technical, and legal base must be carefully developed. We understand the appeal of the approach from a cost-reduction standpoint, but pursuing minimum cost to the exclusion of all other aspects of census quality in 2030 would be akin to weighting confidentiality protection above and beyond all other considerations in 2020. Just as the Census Bureau is still grappling with the anticipated and unanticipated difficulties in completely supplanting its disclosure-avoidance and data-production routines, we worry that rushing to low-cost original enumeration via administrative records raises a great many conceptual and procedural issues that are unlikely to

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

be resolved by 2030, and thus threatens to erode the role of the decennial census as an essential civic ceremony.

12.3.2 Continued Development of Master Address File Curation and Updating

The MAF does play importantly into Recommendation 12.1, from the standpoint of making the MAF a resource covering all living quarters—housing units and GQs alike. But it is also important that several technical aspects of MAF curation and updating are important to address in the coming years, along with clarifying the rules to make participant-input address-building operations more of an active partnership with the Census Bureau.

Recommendation 5.1: The U.S. Census Bureau should continue to research and refine the filters it applies to the Master Address File to derive functional operational extracts, with the intent to reduce the number of addresses cancelled during collection (i.e., flagged as deletes in Nonresponse Followup and other field operations). Along those lines, the Census Bureau should continue to study ways to partner with and more fully utilize U.S. Postal Service data (including the Delivery Sequence File) and potential sources for address addition and revision information between censuses, such as address data recoverable from administrative records extracts, with particular eye toward more regular updating in areas without mail delivery.

Recommendation 5.2: The rules and scope of participant-input address-building operations of the census require clarification, including the Local Update of Census Addresses, New Construction, and Count Review. The U.S. Census Bureau should continue to improve ways to work with user-supplied input or data resources, and make it easier for state, local, and tribal authorities to supply input in usable form, while more clearly laying out the expectations for those address-building operations.

12.3.3 Improvements to Major Census Coverage Evaluation Methods

The 2020 Census experience spotlighted stresses and strains not just on the mechanics of conducting the census but on the primary tools used to assess census quality as well. These are Demographic Analysis, which produces basic counts from sources including birth, death, Medicare, and international migration records and estimates, and comparison of census data with the

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

independent PES. Chapter 4 delves into considerable detail on important improvements that can be made to the application of these evaluation methods in 2030.

Recommendation 4.1: For the Demographic Analysis (DA) coverage evaluation method, the U.S. Census Bureau should conduct research on:

- **The reasonableness of the assumption for incompleteness of Medicare enrollment in the high series, which produced coverage estimates that diverged greatly from the low and middle series for people ages 75 and older, through an appropriate match study;**
- **The suitability of Medicare data for coverage evaluation of the entire population ages 65 and older as in the 2010 Census and for improving the census count of this age group in 2030;**
- **The reasonableness of the methods for assigning Hispanic and non-Hispanic births for coverage estimates for children and young adults, which produced wide differences between the low and high DA series in 2020;**
- **Methods for narrowing estimates of net international migration, which affect estimates for Hispanic people and the total population; and**
- **Methods for developing experimental subnational DA estimates, starting with young children.**

Recommendation 4.3: For the Post-Enumeration Survey (PES) method, the U.S. Census Bureau should:

- **Perform sensitivity analyses, based on plausible assumptions, to put error bounds around such key operations as matching and imputation of match and enumeration status, to evaluate the quality of the 2020 PES and plan such analyses from the outset for the 2030 PES;**
- **Plan analyses for the 2030 PES of components of error for census operations in exhaustive and mutually exclusive categories that are as comparable as possible between 2020 and 2030;**
- **Seek adequate funding to increase the PES sample size in 2030 to at least 2010 levels; and**
- **Experiment with modeling to estimate net undercoverage and overcoverage for more detailed geographic and demographic groups for which direct estimates could not be provided in 2020.**

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Recommendation 4.4: The U.S. Census Bureau should prioritize research on potential sources of coverage errors—both undercounts and overcounts—for geographic areas and population groups, using additional methods besides further analysis of the Post-Enumeration Survey. To address omissions from the census, the Census Bureau should match 2020 Census records with the 2019–2021 American Community Survey and a wide variety of administrative records for census tracts—perhaps sampling those with low self-response rates. The goal would be to provide an evidence base for methods and data sources that could potentially reduce omissions in the 2030 Census. To address duplications in the census, the Census Bureau should match 2020 Census records with themselves to identify duplicate enumerations of people with more than one residence, college students enumerated at college and at home, children in joint custody arrangements, and the like. The goal would be to provide an evidence base for methods to reduce duplicates in the 2030 Census and potentially to remove them from the count, just as people duplicated within a household are dropped from the count.

12.3.4 Revisit Unplanned Contingency Changes from 2020 that Merit Further Consideration

Finally, in taking stock of the procedural history and lessons learned from the 2020 Census, we encourage the Census Bureau to review the operational revisions made dynamically in 2020—in reaction to the COVID-19 pandemic and other difficult enumeration conditions—to consider learnings that could be useful on a planned basis for the 2030 Census. One such unplanned contingency measure that merits consideration for 2030 is the Census Bureau's utilization of its separate Contact Frame to derive phone numbers, which permitted enumerators to make telephone contact with some NRFU cases, particularly in hard-to-access settings. A role for some form of centralized telephone callout as an alternative response mode in NRFU might serve to increase the number of contact attempts made with actual household members and reduce the need for proxy enumeration. Another 2020 contingency measure worthy of assessment and reconsideration is the Every Door Direct Mail postcard that was sent to post office box addresses; it is worthwhile to consider the benefit of a bolstered call to action sent to *all* addresses, not just those with regular delivery by postal mail carriers.

Copyright National Academy of Sciences. All rights reserved.

12.4 IMPROVING CENSUS LAW

We close with a note that is not traditionally a focus of National Academies panels, and hence not an area on which we suggest specific language. But the 2020 Census experience suggests some foundational risks and challenges, for which the only viable solution may be serious reinterpretation and reexamination of census law in Title 13 of the U.S. Code, if not legislative revision of related statutes.

The 400-plus sections of Title 13 of the U.S. Code represent an omnibus compilation of provisions, establishing the Census Bureau, authorizing the Secretary of Commerce to administer various census and statistical data collections, both demographic and economic, and providing for data sharing collaborations among various federal agencies, and in several cases between the federal governments and state and local governments. Title 13 includes quite specific provisions on details ranging from the development and purchase of "mechanical and electronic" equipment to the penalties for violating provisions of the title (including violations of the confidentiality of the information collected). As described in Box 12.1, major changes to Title 13 have been relatively infrequent since the act was codified in 1954.

The crux of the debate over the implementation of new disclosure avoidance methods in the 2020 Census is that the privacy and confidentiality protections of 13 U.S.C. § 9—some version of which have been in statute since 1909—are strong but open to interpretation. It is unclear, for instance, whether the section's prohibition of "any publication whereby the data furnished by any particular establishment or individual under this title can be identified" extends to any characteristics of individuals that might, in combination with other and auxiliary data, leak some personal information. Moreover, Section 9's protections do not operate in a vacuum; there is a wide array of case law and other privacy/confidentiality law alongside which the decennial census and its confidentiality protections must operate. Prominent among these, as discussed in Chapter 9, is the major problem experienced with college and university student housing in the 2020 Census: the Family Educational Rights and Privacy Act's limitation to "directory information" as being the only shareable information, which posed a major challenge as the Census Bureau scrambled to try to get colleges and universities to respond electronically due to campus shutdowns.

In this report, we offer one specific suggestion along these lines—shifting some of the burden of confidentiality protection to the user community.

> **Recommendation 11.5: The U.S. Census Bureau should welcome initiatives to add language to appropriate legislative vehicles, such as the Foundations for Evidence-Based Policymaking Act of 2018, that prescribes responsibilities and penalties for**

Copyright National Academy of Sciences. All rights reserved.

Box 12.1 Major Changes to U.S. Code Title 13 and Census Law Since Codification of Title 13 in 1954

- 1957—authority to use sampling: "Except for the determination of population for apportionment purposes, the Secretary may, where he deems it appropriate, authorize the use of the statistical method known as 'sampling' in carrying out the provisions of this title" (71 Stat. 484; amended in 1976 to replace "may" with "shall" and "apportionment purposes" with "apportionment of Representatives in Congress among the several States" to create current 13 U.S.C. § 195).
- 1964—authority for the mail census: P.L. 88-530 struck requirement that "every enumerator shall visit personally each dwelling house in his subdivision" to collect information, enabling a primarily mail-based census.
- 1975—redistricting data: P.L. 94-171 provided for the reporting of a population tabulation for state legislative apportionment and districting, with the requirement that said data be reported within one year of the census date.
- 1976—authority for mid-decade census: Package of amendments in P.L. 94-521 included authority to conduct a mid-decade census in 1985 and every 10 years thereafter.
- 1994—authority for U.S. Postal Service data sharing and Local Update of Census Addresses program: P.L. 103-340, the Address List Improvement Act of 1994, enabled two major update mechanisms for the current Master Address File, directly authorizing the sharing of address-related information from the U.S. Postal Service (setting up the regular updates from the Delivery Sequence File) and directing the creation of a program by which state and local government liaisons could review and update address information for their jurisdictions.
- 2004—mandate to include "some other race" category on decennial census race question: Provision in appropriations language (not direct amendment to Title 13) barred expenditure of funds for any census or survey not including a "some other race" option, subsequently repeated in several subsequent appropriations bills.

data users, in addition to agency staff, for willful, harmful disclosure of confidential information.

More generally, there may be a need to revisit the interpretation of Title 13's confidentiality protections and reconcile them with provisions in other laws, lest divergent interpretations raise new challenges to census operations.

We also note that the current language of Title 13 does not include an explicit mission statement for the Census Bureau or the decennial census, nor does it speak to the broader purposes of census data collections and products (save for the specificity of the redistricting data file in P.L. 94-171). It could be useful to the vital, participatory civic ceremony that is the U.S. decennial census to have its importance noted—and for the Census Bureau's function as a scientific, statistical agency to be affirmed—in positive law, for 2030 and beyond.

In 2012–2018, it was impossible to know the full array of challenges that would ultimately confront the 2020 Census—and the same can be said looking ahead to the 2030 Census and beyond. There are broader societal and physical

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

trends that portend new, great challenges to the collection of quality census data, among them increasing fluidity of sex and gender identity and increased vulnerability to major disruption from natural (and manmade) disasters. Yet there are also broader forces that may actually be beneficial to the cause of quality census taking, such as wider availability of broadband internet and fuller saturation of smartphone use (and hence comfort with electronic response). Reshaped workforces and workplaces coming out of the COVID-19 pandemic, perhaps with wider adoption of telework and corresponding shifts in the understanding of "usual residence" are also major forces whose effects on census conduct remain to be determined. In this context, it is critically important that the goals and designs for the 2030 Census should be developed in true partnership with census data users and the myriad community of stakeholders and state, local, tribal, and federal government partners—communities defined as broadly as possible—that make the census the vital, grand civic ceremony upon which the nation relies.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

References

2020 Census Quality Indicators Task Force (2020, October). *2020 Census Quality Indicators: A Report from the American Statistical Association*. Alexandria, VA: American Statistical Association.

2020 Census Quality Indicators Task Force (2021, September). *2020 Census State Population Totals: A Report from the American Statistical Association Task Force on 2020 Census Quality Indicators*. Alexandria, VA: American Statistical Association.

Abowd, J., R. Ashmead, R. Cumings-Menon, S. Garfinkel, M. Heineck, C. Heiss, R. Johns, D. Kifer, P. Leclerc, A. Machanavajjhala, B. Moran, W. Sexton, M. Spence, and P. Zhuravlev (2022, June 24). The 2020 Census Disclosure Avoidance System TopDown Algorithm. *Harvard Data Science Review* (Special Issue 2). https://hdsr.mitpress.mit.edu/pub/7evz361i.

Abowd, J. M. (2016, September 15). The challenge of scientific reproducibility and privacy protection for statistical agencies. Presentation to Census Scientific Advisory Committee. https://www2.census.gov/cac/sac/meetings/2016-09/2016-abowd.pdf.

Abowd, J. M. (2018, December 6). Disclosure avoidance for block level data and protection of confidentiality in public tabulations. Presentation to Census Scientific Advisory Committee. https://www2.census.gov/cac/sac/meetings/2018-12/abowd-disclosure-avoidance.pdf.

Abowd, J. M. (2019a, October 18). The U.S. Census Bureau tries to be a good data steward in the 21st century. Presented at FDIC Consumer Research Symposium.

Abowd, J. M. (2019b, January 5). Why the economics profession cannot cede the discussion of privacy protection to computer scientists. Presented at Allied Social Science Associations Annual Meetings.

Abowd, J. M. (2021a, April 13). Declaration. See *State of Alabama v. U.S. Department of Commerce* (2021). Document 41-1.

Copyright National Academy of Sciences. All rights reserved.

Abowd, J. M. (2021b). Official statistics at the crossroads: Data quality and access in an era of heightened privacy risk. *The Survey Statistician* (83), 23–26.

Abowd, J. M. (2021c, April 26). Supplemental declaration. See *State of Alabama v. U.S. Department of Commerce* (2021). Document 116-1.

Abowd, J. M. and M. B. Hawes (2023). Confidentiality protection in the 2020 US Census of Population and Housing. *Annual Review of Statistics and Its Application 10*(1), 119–144.

Abowd, J. M. and I. M. Schmutte (2019, January). An economic analysis of privacy protection and statistical accuracy as social choices. *American Economic Review 109*(1), 171–202.

Advisory Committee on Data for Evidence Building (2022, October 14). *Advisory Committee on Data for Evidence Building: Year 2 Report*. Washington, DC: Advisory Committee on Data for Evidence Building.

Alvarez, J. and E. Maurer (2022, June 21). DHC age data's fitness for use in New York City demographic applications. Presentation at Workshop on the Demographic and Housing Characteristics Files, Washington, DC. https://www.nationalacademies.org/event/06-21-2022/2020-census-data-products-workshop-on-the-demographic-and-housing-characteristics-files.

Alvey, W. and F. Scheuren (1982). Background for an administrative records census. In *Proceedings of the Social Statistics Section, American Statistical Association*, pp. 137–152. Alexandria, VA: American Statistical Association.

American Redistricting Project (2022, June 13). 2020 Census count errors and Congressional apportionment. Blog post: https://thearp.org/blog/apportionment/2020-census-count-errors/.

Antal, L., T. Enderle, and S. Giessing (2017, May 19). *Harmonised protection of census data in the ESS: Statistical disclosure control methods for harmonised protection of census data*. The Hague: Eurostat Centre of Excellence on Statistical Disclosure Control.

Asquith, B. J., B. Hershbein, S. Reed, and S. Yesiltepe (2019, October 25). Differential privacy and scrubbed segregation. Presentation at *Harvard Data Science Review* Conference on Differential Privacy and the 2020 Census.

Author unknown (2015, December 26). Age heaping in the U.S. census. https://journeys.dartmouth.edu/censushistory/2015/12/26/age-heaping-in-the-u-s-census/.

Bach, F. (2022, June). Differential privacy and noisy confidentiality concepts for European population statistics. *Journal of Survey Statistics and Methodology* (3), 642–687.

Bailie, J., R. Gong, and X.-L. Meng (2023, May). Can swapping be differentially private? a refreshment stirred, not shaken. Technical report, National Bureau of Economic Research, https://conference.nber.org/conf_papers/f178188.pdf.

Copyright National Academy of Sciences. All rights reserved.

Beaghen, M., R. N. Turner, M. Jost, and E. Jarra (2022, May 19). *2020 Post-Enumeration Survey Estimation Methods: Missing Data for Person Estimates*. DSSD 2020 Post-Enumeration Survey Memorandum Series #2020-J-05. Washington, DC: U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/coverage-measurement/pes/estimation-methods-addressing-missing-data-for-people.pdf.

Biemer, P., J. Salvo, and J. Auerbach (2021, September). *The Quality of the 2020 Census: An Independent Assessment of Census Bureau Activities Critical to Data Quality*. First submitted to 2020 Census Quality Indicators Task Force July 22, 2021; revised August 24, 2021, and released alongside task force report. Alexandria, VA: American Statistical Association.

Bowlby, G. and M. Beaulieu (2023, May 24). International approaches to conducting a census in the 21st century: Canada. Presentations in Committee on National Statistics Public Webinar Series. https://www.nationalacademies.org/event/05-24-2023/cnstat-public-webinar-series-international-approaches-to-conducting-a-census-in-the-21st-century-canada.

Brown, J. D., S. R. Cohen, G. Denoeux, S. Dorinski, M. L. Heggeness, C. Lieberman, L. McBride, M. Murray-Close, H. Qin, A. E. Ross, D. H. Sandler, L. Warren, and M. Yi (2023, May 5). *Real-Time 2020 Administrative Record Census Simulation* (1.0 ed.). 2020 Census Experiment Report, released via 2020 Census Program Memorandum Series #2023.08. Washington, DC: U.S. Census Bureau.

Brummet, Q., E. Mulrow, and K. Wolter (2019, October 25). The effect of differentially private noise injection on sampling efficiency and funding allocations: Evidence from the 1940 Census. Presentation at *Harvard Data Science Review* Conference on Differential Privacy and the 2020 Census.

Bycroft, C. and N. Matheson-Dunning (2023, June 7). International approaches to conducting a census in the 21st century: New Zealand. Presentations in Committee on National Statistics Public Webinar Series. https://www.nationalacademies.org/event/06-07-2023/cnstat-public-webinar-series-international-approaches-to-conducting-a-census-in-the-21st-century-new-zealand.

Cassal, K. R. (2022, June 21). Making DHC analysis more accessible and exploring solutions for data users. Presentation at Workshop on the Demographic and Housing Characteristics Files, Washington, DC. https://www.nationalacademies.org/event/06-21-2022/2020-census-data-products-workshop-on-the-demographic-and-housing-characteristics-files.

Census Geographies Project (2022, December). *U.S. Political and Statistical Geography and Census Small-Area Statistics: A Primer*. Washington, DC: Massive Data Institute, McCourt School of Public Policy, Georgetown University.

Citro, C. F. (2011). Editing and imputation. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census: From the Constitution to the American Community Survey*. Washington, DC: CQ Press.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Citro, C. F. (2021, October 16). Preliminary assessment of aspects of census quality via demographic analysis. https://www.documentcloud.org/documents/21097003-constance-citro-preliminary-assessment-of-aspect-of-2020-census-quality-via-demographic-analysis-oct-16-2021-version.

Craigle, M. (2022, June 21). How decennial data drive funding at state and local levels—some examples from Montana. Presentation at Workshop on the Demographic and Housing Characteristics Files, Washington, DC. https://www.nationalacademies.org/event/06-21-2022/2020-census-data-products-workshop-on-the-demographic-and-housing-characteristics-files.

Cresce, A. R. and R. R. Ramirez (2003). *Analysis of General Hispanic Responses in Census 2000*. Population Division Working Paper No. 72. Washington, DC: U.S. Census Bureau. https://www2.census.gov/library/working-papers/2003/demographic/pop-twps0072/pop-twps0072.pdf.

Daily, D. (2022, December 14). Disclosure avoidance protections for the American Community Survey. Random Samplings blog. https://www.census.gov/newsroom/blogs/random-samplings/2022/12/disclosure-avoidance-protections-acs.html.

Dajani, A. N., A. D. Lauger, P. E. Singer, D. Kifer, J. P. Reiter, A. Machanavajjhala, S. L. Garfinkel, S. A. Dahl, M. Graham, V. Karwa, H. Kim, P. Leclerc, I. M. Schmutte, W. N. Sexton, L. Vilhuber, and J. M. Abowd (2017, September 21). The modernization of statistical disclosure limitation at the U.S. Census Bureau. Presentation at Joint UNECE/EUROSTAT Work Session on Statistical Data Confidentiality, Skopje. https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2017/3_Dajani.pdf.

Davis, P. P. and J. Mulligan (2012, May 22). *Net coverage for the household population in the United States*. DSSD 2010 Census Coverage Measurement Memorandum Series No. 2010-G-03. Washington, DC: U.S. Department of Commerce.

Decennial Census Management Division (2018, December 31). *2020 Census Detailed Operational Plan for: 6. Geographic Programs Operation (GEOP)—6-1. Geographic Delineations Component (GEOP/GD)*. Washington, DC: U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/2020/program-management/planning-docs/GEOP_GD_detailed_operational_plan_v2.pdf.

DeJesus, J. and S. Konya (2023, February 7). *2020 Census Item Nonresponse and Imputation Assessment Report*. Version 1.0. Washington, DC: U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/2020/program-management/evaluate-docs/EAE-2020-item-nonresponse-assessment.pdf.

Department of Commerce v. New York (2019). 588 U.S. ___, 139 S. Ct. 2551. No. 18-966. Argued April 23, 2019. Decided June 27, 2019.

Devine, J., R. Bhaskar, B. DeSalvo, J. G. Robinson, M. Scopilliti, and K. K. West (2012, February). *The Development and Sensitivity Analysis of the 2010 Demographic Analysis Estimates*. Population Division Working Paper No. 93. Washington, DC: U.S.

Copyright National Academy of Sciences. All rights reserved.

Census Bureau. https://www.census.gov/content/dam/Census/library/working-papers/2012/demo/POP-twps0093.pdf.

Devine, J., M. Spence, A. Krause, R. Ramirez, and M. Hawes (2023, February 8). 2020 Census data products workshop [presentation slides]. Presented at Applied Demography Conference, Annapolis, MD.

Donello, D. (2021). Overview of Non-ID processing. U.S. Census Bureau. Presentation to the Panel to Evaluate the Quality of the 2020 Census, August 30, 2021, virtual format.

Dove, I. (2021, August 20). *Applying Differential Privacy Protection to ONS Mortality Data, Pilot Study*. Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/methodologies/applyingdifferentialprivacyprotectiontoonsmortalitydatapilotstudy.

Efron, B. (1978). Regression and ANOVA with zero-one data: Measures of residual variation. *Journal of the American Statistical Association 73*(361), 113–121.

Equitable Data Working Group (2022, April 22). A vision for equitable data: Recommendations from the Equitable Data Working Group. https://www.whitehouse.gov/wp-content/uploads/2022/04/eo13985-vision-for-equitable-data.pdf.

Executive Order No. 13880 (2019, July 11). *Collecting Information About Citizenship Status in Connection with the Decennial Census*. 84 FR 33821.

Executive Order No. 13985 (2021, January 20). *Advancing Racial Equity and Support for Underserved Communities Through the Federal Government*. 86 FR 7009.

Fay, R. E. (2021, September). *Review of Research Supporting the Use of Administrative Record Enumeration in the 2020 Census*. Submitted to 2020 Census Quality Indicators Task Force August 12, 2021; released alongside task force report. Alexandria, VA: American Statistical Association.

Federal Committee on Statistical Methodology (2020, September). *A Framework for Data Quality*. FCSM-20-04. Washington, DC: Federal Committee on Statistical Methodology.

Federal State Cooperative on Population Estimates Steering Committee (2020, November 23). Letter to members of the U.S. Census Bureau Data Stewardship Executive Policy Committee. https://ofm.wa.gov/sites/default/files/public/dataresearch/pop/sdc/FSCPE_letter_to_DSEP_re_determination_of_invariants.pdf.

Federal State Cooperative on Population Estimates Steering Committee (2022, August 1). Letter to Robert Santos, director, U.S. Census Bureau [with additional signatories]. https://pad.human.cornell.edu/papers/downloads/FSCPE-SC-LetterToDirectorSantos-8-22-22.pdf.

Fontenot, A. E. (2018a, May 16). *2020 Census Plans for the Commonwealth of Puerto Rico*. 2020 Census Program Memorandum Series #2018.13. Washington, DC: U.S. Census Bureau.

Copyright National Academy of Sciences. All rights reserved.

Fontenot, A. E. (2018b, February 15). *2020 Census Residence Criteria and Residence Situations*. 2020 Census Program Memorandum Series #2018.04. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. (2018c, January 26). *Using Two Separate Questions for Race and Ethnicity in 2018 End-to-End Census Test and 2020 Census*. 2020 Census Program Memorandum Series #2018.02. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. (2019, July 1). *Disclosure Avoidance System Parameters and Global Privacy-Loss Budget for the 2018 End-to-End Census Test*. 2020 Census Program Memorandum Series #2019.13. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. (2021a, November 15). *Documentation of Changes to the Count Review Operation*. 2020 Census Program Memorandum Series #2021.22. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. (2021b, April 6). *Documentation of Updates to the Enumeration at Transitory Locations Operation*. 2020 Census Program Memorandum Series #2021.08. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. (2021c, February 12). *Documentation of Updates to the Non-ID Processing Operation*. 2020 Census Program Memorandum Series #2021.03. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. (2021d, October 18). *Documentation of Updates to the Nonresponse Followup Operation*. 2020 Census Program Memorandum Series #2021.20. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. (2021e, January 21). *Documention of Updates to the Internet Self-Response Operation*. 2020 Census Program Memorandum Series #2021.04. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. (2021f, January 25). *Documention of Updates to the Update Leave Operation*. 2020 Census Program Memorandum Series #2021.05. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. (2021g, September 9). *Group Quarters Operation Detailed Operational Plan v2.0*. 2020 Census Program Memorandum Series #2021.16. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. (2021h). Overview of operations and challenges in the 2020 census. U.S. Census Bureau. Presentation to the Panel to Evaluate the Quality of the 2020 Census, July 19, 2021, virtual format.

Fontenot, A. E. (2021i, March 2). *Updates to the 2020 Update Enumerate Detailed Operational Plan*. 2020 Census Program Memorandum Series #2021.06. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. (2021j, January 12). *Updates to the Census Questionnaire Assistance Operation Scope*. 2020 Census Program Memorandum Series #2021.02. Washington, DC: U.S. Census Bureau.

Copyright National Academy of Sciences. All rights reserved.

Fontenot, A. E. (2021k, September 28). *Updates to the Evaluations and Experiments Operation*. 2020 Census Program Memorandum Series #2021.17. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. (2022a, March 3). *Census 2020 High-Level Overview of Group Quarters Count Imputation*. 2020 Census Program Memorandum Series #2022.08. Washington, DC: U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/2020/program-management/memo-series/2020-memo-2022_08.pdf.

Fontenot, A. E. (2022b, January 6). *Updates to the 2020 Census Response Processing Operation Detailed Operational Plan*. 2020 Census Program Memorandum Series #2022.01. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. (2022c, January 10). *Updates to the Geographic Programs Operation, Geographic Delineations Detailed Operational Plan*. 2020 Census Program Memorandum Series #2022.02. Washington, DC: U.S. Census Bureau.

Fontenot, A. E. and B. Taylor (2021). Census budget and possible impact of budget reductions. U.S. Census Bureau. Presentation to the Panel to Evaluate the Quality of the 2020 Census, September 28, 2021, virtual format.

Garfinkel, S. (2023). Comment to Mulalidhar and Domingo-Ferrer (2023)—legacy statistical disclosure limitation techniques were not an option for the 2020 US Census of Population and Housing. *Journal of Official Statistics 39*(3), 399–410. https://doi.org/10.2478/jos-2023-0018.

Garfinkel, S. L. (2017, September 15). Modernizing disclosure avoidance: Report on the 2020 disclosure avoidance subsystem as implemented for the 2018 end-to-end test (continued). Presentation to Census Scientific Advisory Committee. https://www2.census.gov/cac/sac/meetings/2016-09/2016-abowd.pdf.

Garner, E. (2023, April). The use of differential privacy by the U.S. Census Bureau—a state's perspective. Presented at Population Association of America meetings, Atlanta, GA.

Gibson, C. J. and K. Jung (2002, September). Historical census statistics on population totals by race, 1790 to 1990, and by Hispanic origin, 1970 to 1990, for the United States, regions, divisions, and states. Working Paper 56, U.S. Census Bureau, Population Division.

Gong, R., E. L. Groshen, and S. Vadhan (2022, June 24). Harnessing the Known Unknowns: Differential Privacy and the 2020 Census. *Harvard Data Science Review* (Special Issue 2). https://hdsr.mitpress.mit.edu/pub/fgyf5cne.

Grieco, E. M. and R. C. Cassidy (2001, March). *Overview of Race and Hispanic Origin, 2000*. Census 2000 Brief C2KBR/01-1. Washington, DC: U.S. Census Bureau.

Groshen, E. L. and D. Goroff (2022, June 24). The 2020 Census Disclosure Avoidance System TopDown Algorithm. *Harvard Data Science Review* (Special Issue 2). https://hdsr.mitpress.mit.edu/pub/tgi3gol2/release/2.

Copyright National Academy of Sciences. All rights reserved.

Hardcastle, J. (2019, December 12). Housing and population counts—implications for local estimates and projections. Presentation at Workshop on 2020 Census Data Products: Data Needs and Privacy Considerations, Washington, DC. https://www.nationalacademies.org/event/12-11-2019/workshop-on-2020-census-data-products-data-needs-and-privacy-considerations.

Harding, S., L. Liddle, P. McDonald, P. Morrison, D. Trewin, and S. Walters (2022, June 28). Report on the quality of 2021 Census data. Report of the Statistical Independent Assurance Panel to the Chief Statistician, Australian Bureau of Statistics. https://www.abs.gov.au/census/about-census/census-statistical-independent-assurance-panel-report.

Hartley, C. (2022, June 21). Status update on the population estimates. Presentation at Workshop on the Demographic and Housing Characteristics Files, Washington, DC. https://www.nationalacademies.org/event/06-21-2022/2020-census-data-products-workshop-on-the-demographic-and-housing-characteristics-files.

Hearns, K. (2021). Overview of internet self-response instrument and contact strategies development. U.S. Census Bureau. Presentation to the Panel to Evaluate the Quality of the 2020 Census, August 30, 2021, virtual format.

Heim, K. (2022, September 21). *2020 Post-Enumeration Survey Estimation Methods: Net Coverage Estimation*. DSSD 2020 Post-Enumeration Survey Memorandum Series #2020-J-07. Washington, DC: U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/coverage-measurement/pes/estimation-methods-net-coverage-estimation.pdf.

Henry, S., R. Pereira, and L. Blackwell (2023, May 18). International approaches to conducting a census in the 21st century: United Kingdom. Presentations in Committee on National Statistics Public Webinar Series. https://www.nationalacademies.org/event/05-18-2023/cnstat-public-webinar-series-international-approaches-to-conducting-a-census-in-the-21st-century-uk.

Herdağdelen, A., A. Dow, B. State, P. Mohassel, and A. Pompe (2020, June 2). Protecting privacy in Facebook mobility data during the COVID-19 response. https://research.facebook.com/blog/2020/06/protecting-privacy-in-facebook-mobility-data-during-the-covid-19-response.

Hill, C., K. Heim, J. Hong, and N. Phan (2022, June). *Census Coverage Estimates for People in the United States by State and Census Operations: 2020 Post-Enumeration Survey Estimation Report*. PES20-G-02RV. Washington, DC: U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/coverage-measurement/pes/census-coverage-estimates-for-people-in-the-united-states-by-state-and-census-operations.pdf.

Hobbs, F. (2004). Age and sex composition. In J. Siegel and D. Swanson (Eds.), *The Method and Materials of Demography* (Second ed.)., pp. 125–173. Elsevier Academic Press.

Copyright National Academy of Sciences. All rights reserved.

Hodges, K. and R. Cortes (2022, June 21). Population and householders by age before and after differential privacy. Presentation at Workshop on the Demographic and Housing Characteristics Files, Washington, DC. https://www.nationalacademies.org/event/06-21-2022/2020-census-data-products-workshop-on-the-demographic-and-housing-characteristics-files.

Holmberg, A. and R. Watmuff (2023, May 31). International approaches to conducting a census in the 21st century: Australia. Presentations in Committee on National Statistics Public Webinar Series. https://www.nationalacademies.org/event/05-31-2023/cnstat-public-webinar-series-international-approaches-to-conducting-a-census-in-the-21st-century-australia.

Hotz, V. J., C. R. Bollinger, T. Komorova, and B. D. Spencer (2022, July 25). Balancing data privacy and usability in the federal statistical system. *Proceedings of the National Academy of Sciences 119*(31), e2104906119.

Hotz, V. J. and J. Salvo (2022, June 24). A chronicle of the application of differential privacy to the 2020 Census. *Harvard Data Science Review* (Special Issue 2). https://hdsr.mitpress.mit.edu/pub/7evz361i.

Jackson, G., K. Wechter, and S. Winder (2012, May 22). *2010 Census Be Counted and Questionnaire Assistance Centers Assessment*. 2010 Census Planning Memoranda Series No. 194. Washington, DC: U.S. Census Bureau.

Jarmin, R. (2019a, February 15). Census Bureau adopts cutting edge privacy protections for 2020 Census. Random Samplings blog. https://www.census.gov/newsroom/blogs/random-samplings/2019/02/census_bureau_adopts.html.

Jarmin, R. (2019b, July 30). Census Bureau continues to boost data safeguards. Random Samplings blog. https://www.census.gov/newsroom/blogs/random-samplings/2019/07/boost-safeguards.html.

JASON (2016, November). *Alternative Futures for the Conduct of the 2030 Census*. McLean, VA: The MITRE Corporation. JSR-16-Task-009. https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/alternative-futures-2030-census.pdf.

JASON (2021, February 8). *JSR-20-2N (WS'21) Assessment of 2020 Census Data Quality Processes*. McLean, VA: The MITRE Corporation. Letter report to Christa Jones and Deborah Stempowski, U.S. Census Bureau.

Jensen, E. and T. Kennel (2022, March 10). Who was undercounted, overcounted in the 2020 Census? Detailed coverage estimates for the 2020 Census release today. https://www.census.gov/library/stories/2022/03/who-was-undercounted-overcounted-in-2020-census.html.

Jensen, E., A. Roberts, and L. Rogers (2023, May 25). Age heaping in the 2020 Census Demographic and Housing Characteristics File (DHC). Posted to Random Samplings blog. https://www.census.gov/newsroom/blogs/random-samplings/2023/05/age-heaping-2020-census-dhc.html.

Copyright National Academy of Sciences. All rights reserved.

Jensen, E. B., A. Knapp, H. King, D. Armstrong, S. L. Johnson, L. Sink, and E. Miller (2020, December 15). Methodology for the 2020 demographic analysis estimates. Technical report, U.S. Census Bureau, Population Division.

Johanson, C., M. Scheu, and K. Wechter (2011, December 19). *2010 Census Operational Assessment for Type of Enumeration Area Delineation*. 2010 Census Planning Memoranda Series No. 164. Washington, DC: U.S. Census Bureau.

Jones, N., R. Ramirez, and R. Marks (2021). Improvements to race and ethnicity coding in the 2020 Census. U.S. Census Bureau. Presentation to the Panel to Evaluate the Quality of the 2020 Census, September 28, 2021, virtual format.

Jost, M. and S. Khubba (2022). *2020 Post-Enumeration Survey: Net Coverage Error and Components of Census Coverage for Race Groups and Hispanic Origin by Age, Sex, and Tenure in the United States*. DSSD 2020 Census Post-Enumeration Survey Memorandum Series No. 2010-G-05. Washington, DC: U.S. Department of Commerce.

Kaneff, J. (2022, June 21). Regional and local government planning: Decennial census data needs and assessment. Presentation at Workshop on the Demographic and Housing Characteristics Files, Washington, DC. https://www.nationalacademies.org/event/06-21-2022/2020-census-data-products-workshop-on-the-demographic-and-housing-characteristics-files.

Keller, A. and T. Fox (2012, May 22). *Components of census coverage for the household population in the United States*. DSSD 2010 Census Coverage Measurement Memorandum Series No. 2010-G-04. Washington, DC: U.S. Department of Commerce.

Keller, A. and T. Fox (2013, May 23). *2010 Components of Census Coverage for Race Groups and Hispanic Origin by Age, Sex, and Tenure in the United States*. DSSD 2010 Census Coverage Measurement Memorandum Series No. 2010-E-51. Publicly released March 2023. Washington, DC: U.S. Department of Commerce.

Keller, S. and S. Shipp (2021, February 4). A new approach for the 2030 decennial census; a 21st century census for a 21st century society. University of Virginia Biocomplexity Institute White Paper. https://biocomplexity.virginia.edu/sites/default/files/census/white-papers/2030Census%20April%202022.pdf.

Keller, S. A. and J. M. Abowd (2023). Database reconstruction does compromise confidentiality. *PNAS 120*(12), 1–3. https://doi.org/10.1073/pnas.2300976120.

Kenny, C. T., S. Kuriwaki, C. McCartan, E. T. Rosenman, T. Simko, and K. Imai (2021, October 6). The use of differential privacy for census data and its impact on redistricting: The case of the 2020 U.S. Census. *Science Advances 7*(41), eabk3283.

Kenny, C. T., S. Kuriwaki, C. McCartan, T. Simko, and K. Imai (2023, June 13). Evaluating bias and noise induced by the U.S. Census Bureau's privacy protection methods. arXiv:2306.07521v1.

Copyright National Academy of Sciences. All rights reserved.

Khubba, S., K. Heim, and J. Hong (2022, March). *National Census Coverage Estimates for People in the United States by Demographic Characteristics: 2020 Post-Enumeration Survey Estimation Report*. PES20-G-01. Washington, DC: U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/coverage-measurement/pes/national-census-coverage-estimates-by-demographic-characteristics.pdf.

Lane, L., S. Hanks, A. Williams, and K. Roinestad (2022, August 18). *2020 Census In-Field Address Canvassing Operational Assessment Report*. Washington, DC: U.S. Census Bureau.

Layne, M., D. Wagner, and C. Rothhaas (2014). Estimating record linkage false match rate for the Person Identification Validation System. Center for Administrative Records Research and Applications Working Paper 2014-02, U.S. Census Bureau, Washington, DC.

Letourneau, E. (2012, June 6). *Mail Response/Return Rates Assessment*. 2010 Census Planning Memoranda Series No. 198. Washington, DC: U.S. Census Bureau.

Liebler, C. A. (2022, June 21). Impacts of differential privacy on data about American Indian and Alaska Native people and tribal lands. Presentation at Workshop on the Demographic and Housing Characteristics Files, Washington, DC. https://www.nationalacademies.org/event/06-21-2022/2020-census-data-products-workshop-on-the-demographic-and-housing-characteristics-files.

Lyons-Amos, M. and T. Stones (2017). Trends in demographic and health survey data quality: an analysis of age heaping over time in 34 countries in Sub-Saharan Africa between 1987 and 2015. *BMC Res Notes 10*, 760.

Macagnone, M. (2022, April 5). Census Bureau director on 'transformational' changes—and SXSW. https://rollcall.com/2022/04/05/census-bureau-director-on-transformational-changes-and-sxsw/.

Marks, R. and M. Rios-Vargas (2021, August 3). Improvements to the 2020 Census race and hispanic origin question designs, data processing, and coding procedures. Posted to Census Bureau Random Samplings blog. https://www.census.gov/newsroom/blogs/random-samplings/2021/08/improvements-to-2020-census-race-hispanic-origin-question-designs.html.

Marra, E. and T. Kennel (2022, March). *Source and Accuracy of the 2020 Post-Enumeration Survey Person Estimates: 2020 Post-Enumeration Survey Methodology Report*. PES20-J-01. Washington, DC: U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/coverage-measurement/pes/2020-source-and-accuracy-pes-estimates.pdf.

Mathews, K., J. Phelan, N. A. Jones, S. Konya, R. Marks, B. M. Pratt, J. Coombs, and M. Bentley (2017, February 28). *2015 National Content Test: Race and Ethnicity Analysis Report*. Version 1.0. Washington, DC: U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2015nct-race-ethnicity-analysis.pdf.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

McGeeney, K., B. Kriz, S. Mullenax, L. Kail, G. Walejko, M. Vines, N. Bates, and Y. G. Trejo (2019, January 24). *2020 Census Barriers, Attitudes, and Motivators Study Survey Report*. Version 2.0. Washington, DC: U.S. Census Bureau.

McKenna, L. (2018, November). *Disclosure Avoidance Techniques Used for the 1970 through 2010 Decennial Censuses of Population and Housing*. CES-WP-18-47. Washington, DC: U.S. Census Bureau.

Mervis, J. (2017, July 25). Researchers think they've found a much better way to conduct the 2030 U.S. census. https://www.science.org/content/article/researchers-think-they-ve-found-much-better-way-conduct-2030-us-census.

Mueller, J. T. (2022, June 21). County-level discrepancies between SF-1 and DHC across the rural-urban continuum. Presentation at Workshop on the Demographic and Housing Characteristics Files, Washington, DC. https://www.nationalacademies.org/event/06-21-2022/2020-census-data-products-workshop-on-the-demographic-and-housing-characteristics-files.

Mule, T. (2012, May 22). *Summary of estimates of coverage for persons in the United States*. DSSD 2010 Census Coverage Measurement Memorandum Series No. 2010-G-01. Washington, DC: U.S. Department of Commerce.

Mulry, M. H. and B. D. Spencer (1991, December). Total error in PES estimates of population: The dress rehearsal census of 1988. *Journal of the American Statistical Association 86*, 839–854.

Mulry, M. H., C. J. Tello-Trillo, T. Mule, and A. Keller (2023a, February 21). Comparisons of administrative record rosters to census self-responses and NRFU household member responses. Research Report Series 2023-01, U.S. Census Bureau, Washington, DC. https://www2.census.gov/library/working-papers/2023/adrm/RRS2023-01.pdf.

Mulry, M. H., C. J. Tello-Trillo, T. Mule, and A. Keller (2023b, February 27). Full report of the comparisons of administrative record rosters to census self-responses and NRFU household member responses. Working Paper DEC-WP-2023-01, U.S. Census Bureau, Washington, DC. https://www2.census.gov/library/working-papers/2023/decennial/dec-wp-2023-01.pdf.

Muralidhar, K. and J. Domingo-Ferrer (2023a). Database reconstruction is not so easy and is different from reidentification. *Journal of Official Statistics 39*(3), 381–398. https://doi.org/10.2478/jos-2023-0017.

Muralidhar, K. and J. Domingo-Ferrer (2023b). A rejoinder to Garfinkel (2023)—legacy statistical disclosure limitation techniques for protecting 2020 Decennial US Census: Still a viable option. *Journal of Official Statistics 39*(3), 411–420. https://doi.org/10.2478/jos-2023-0019.

Nagle, N. and T. Kuhn (2019, December 11). Implications for municipalities and school enrollment statistics. Presentation at Workshop on 2020

Copyright National Academy of Sciences. All rights reserved.

Census Data Products: Data Needs and Privacy Considerations, Washington, DC. https://www.nationalacademies.org/event/12-11-2019/workshop-on-2020-census-data-products-data-needs-and-privacy-considerations.

National Academies of Sciences, Engineering, and Medicine (2015). *The Bicentennial Census: New Directions for Methodology in 1990: 30th Anniversary Edition*. Washington, DC: The National Academies Press.

National Academies of Sciences, Engineering, and Medicine (2017). *Innovations in Federal Statistics: Combining Data Sources While Protecting Privacy*. Washington, DC: The National Academies Press.

National Academies of Sciences, Engineering, and Medicine (2020). *2020 Census Data Products: Data Needs and Privacy Considerations: Proceedings of a Workshop*. Washington, DC: The National Academies Press.

National Academies of Sciences, Engineering, and Medicine (2021). *Principles and Practices for a Federal Statistical Agency* (7th ed.). Washington, DC: The National Academies Press.

National Academies of Sciences, Engineering, and Medicine (2022). *Understanding the Quality of the 2020 Census: Interim Report*. Washington, DC: The National Academies Press.

National Academies of Sciences, Engineering, and Medicine (2023). *2020 Census Data Products: Demographic and Housing Characteristics File: Proceedings of a Workshop*. Washington, DC: The National Academies Press.

National Conference of State Legislatures (2021, May 26). Redistricting and use of census data. https://www.ncsl.org/redistricting-and-census/redistricting-and-use-of-census-data.

National Research Council (1985). *The Bicentennial Census: New Directions for Methodology in 1990*. Panel on Decennial Census Methodology, Constance F. Citro and Michael L. Cohen, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (1994). *Counting People in the Information Age*. Panel to Evaluate Alternative Census Methods, Duane L. Steffey and Norman M. Bradburn, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (1995). *Modernizing the U.S. Census*. Panel on Census Requirements in the Year 2000 and Beyond, Barry Edmonston and Charles Schultze, eds., Committee on National Statistics, Commission of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

National Research Council (2004a). *The 2000 Census: Counting Under Adversity*. Panel to Review the 2000 Census, Constance F. Citro, Daniel L. Cork, and Janet L. Norwood, eds., Committee on National Statistics. Washington, DC: The National Academies Press.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

National Research Council (2004b). *Reengineering the 2010 Census: Risks and Challenges*. Panel on Research on Future Census Methods, Daniel L. Cork, Michael L. Cohen, and Benjamin F. King, eds., Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council (2006). *Once, Only Once, and in the Right Place: Residence Rules in the Decennial Census*. Panel on Residence Rules in the Decennial Census, Daniel L. Cork and Paul R. Voss, eds., Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council (2008). *Coverage Measurement in the 2010 Census*. Panel on Correlation Bias and Coverage Measurement in the 2010 Decennial Census, Robert M. Bell and Michael L. Cohen, eds., Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council (2009). *Coverage Measurement in the 2010 Census*. Panel on Correlation Bias and Coverage Measurement in the 2010 Decennial Census, Robert M. Bell and Michael L. Cohen, eds., Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council (2010). *Envisioning the 2020 Census*. Panel on the Design of the 2010 Census Program of Evaluations and Experiments, Lawrence D. Brown, Michael L. Cohen, Daniel L. Cork, and Constance F. Citro, eds. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council (2011). *Change and the 2020 Census: Not Whether But How*. Panel to Review the 2010 Census, Thomas M. Cook, Janet L. Norwood, and Daniel L. Cork, eds. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Urban League v. Ross (2020, August 18). Case No. 5:20-cv-05799, N.D. California.

Neunhoeffer, M., D. Sheldon, and A. Smith (2023, May 2). Bootstrap-based, general-purpose statistical inference from differential private releases. Technical report, National Bureau of Economic Research, https://conference.nber.org/conf_papers/f178219.pdf.

O'Hara, A. (2022, June 21). Observations on use cases and needs. Presentation at Workshop on the Demographic and Housing Characteristics Files, Washington, DC. https://www.nationalacademies.org/event/06-21-2022/2020-census-data-products-workshop-on-the-demographic-and-housing-characteristics-files.

O'Hare, W. P. (2021, May 5). Measuring the quality of the 2020 Census: What do we know now? Presented at Population Association of America meetings.

Copyright National Academy of Sciences. All rights reserved.

O'Hare, W. P. and J. J. J. Lee (2021, April). *Who Responded in the 2020 Census? Variation in Tract-Level Response Rates in the 2020 U.S. Census*. Washington, DC: Georgetown Center on Poverty and Inequality.

Pardeshi, G. S. (2010, July). Age heaping and accuracy of age data collected during a community survey in the Yavatmal District, Maharashtra. *Indian Journal of Community Medicine 35*(3), 391–395.

Passel, J. S. and J. M. Krogstad (2023, April 19). How a coding error provided a rare glimpse into Latino identity among Brazilians in the U.S. https://www.pewresearch.org/short-reads/2023/04/19/how-a-coding-error-provided-a-rare-glimpse-into-latino-identity-among-brazilians-in-the-u-s/.

Pew Research Center (2022, June 6). Public trust in government: 1958–2022. https://www.pewresearch.org/politics/2022/06/06/public-trust-in-government-1958-2022/.

Phan, N. and J. Lawrence (2022, May 19). *2020 Post-Enumeration Survey Estimation Methods: Characteristic Editing and Imputation*. DSSD 2020 Post-Enumeration Survey Memorandum Series #2020-J-04. Washington, DC: U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/coverage-measurement/pes/estimation-methods-characteristic-imputation.pdf.

Poston, D. L. and L. F. Bouvier (2017). *Population and Society: An Introduction to Demography* (Second ed.). Cambridge: Cambridge University Press.

Presidential Memorandum for the Secretary of Commerce (2020, July 23). *Excluding Illegal Aliens from the Apportionment Base Following the 2020 Census*. 85 FR 44679.

Pullum, T. W. (2006). *An Assessment of Age and Date reporting in the DHS Surveys, 1985–2003*. U.S. Agency for International Development, DHS Methodological Reports No. 5. Calverton, Maryland: Macro International Inc.

Reamer, A. (2020, April 29). Counting for dollars 2020: The role of the decennial census in the geographic distribution of federal funds. https://gwipp.gwu.edu/counting-dollars-2020-role-decennial-census-geographic-distribution-federal-funds.

Reynolds, L. and J. Vink (2022, June 21). Differential privacy, differential accuracy?: A use case on data segmentation by tract-level tenure majority. Presentation at Workshop on the Demographic and Housing Characteristics Files, Washington, DC. https://www.nationalacademies.org/event/06-21-2022/2020-census-data-products-workshop-on-the-demographic-and-housing-characteristics-files.

Reynolds, L. and J. Vink (2023, February 9). Homeownership rates in a differentially private census: Racial disparities and geographic dependency. Presented at Applied Demography Conference, Annapolis, MD.

Richmond, L. and S. Hanks (2022, August 18). *2020 Census In-Office Address Canvassing Operational Assessment Report*. Washington, DC: U.S. Census Bureau.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Ridgeway, D., M. F. Theofanos, T. W. Manley, and C. Task (2021, April). *Challenge Design and Lessons Learned from the 2018 Differential Privacy Challenges*. NIST Technical Note 2151. Washington, DC: National Institute of Standards and Technology.

Robinson, J. G. (2010, January 8). Coverage of population in Census 2000 based on demographic analysis: The history behind the numbers. Working Paper 91, U.S. Census Bureau, Population Division.

Ruggles, S., C. Fitch, D. Magnuson, and J. Schroeder (2019). Differential privacy and census data: Implications for social and economic research. *AEA Papers and Proceedings 109*(May 2019), 403–408. https://www.aeaweb.org/articles?id=10.1257/pandp.20191107.

Shlomo, N. (2023, June 30). Statistical disclosure control and developments in formal privacy: In memoriam to Chris Skinner. *Survey Methodology 49*(1), 2–25.

Spoorenberg, T. (2009). Is the Whipple's index really a fair and reliable measure of the quality of age reporting? an analysis of 234 censuses from 145 countries. Conference paper. https://ipc2009.popconf.org/papers/90551.

State of Alabama v. U.S. Department of Commerce (2021). Case No. 3:21-cv-00211, M.D. Alabama. Filed March 10, 2021. Dismissal order issued September 9, 2021.

Templ, M. and M. Sariyar (2022, August). A systematic overview on methods to protect sensitive data provided for various analyses. *International Journal of Information Security 21*, 1233–1246.

Thieme, M. (2021, March 9). Finding 'anomalies' illustrates 2020 Census quality checks are working. Published on U.S. Census Bureau Random Samplings blog. https://www.census.gov/newsroom/blogs/random-samplings/2021/03/finding_anomalies.html.

United Nations (2007). *2004 Demographic Yearbook, Fifty-sixth Issue*. ST/ESA/STAT/SER.R/35. New York: United Nations, Department of Economic and Social Affairs.

United Nations (2014, March 3). *Fundamental Principles of Official Statistics*. New York. General Assembly Resolution A/RES/68/261 adopted January 29, 2014. https://unstats.un.org/fpos.

U.S. Census Bureau (1966). *1960 Censuses of Population and Housing: Procedural History*. Washington, DC: U.S. Government Printing Office.

U.S. Census Bureau (1973, June). *1970 Census of Population, Subject Reports: Persons of Spanish Origin*. Report PC(2)-1C. Washington, DC: U.S. Government Printing Office.

U.S. Census Bureau (2001, April 2). *Population by Race and Hispanic or Latino Origin for the United States: 1990 and 2000*. Report PHC-T-1. Washington, DC: U.S. Government Printing Office.

Copyright National Academy of Sciences. All rights reserved.

U.S. Census Bureau (2012a, July). 2010 Summary File 2—2010 Census of Population and Housing—technical documentation. https://www2.census.gov/programs-surveys/decennial/2010/technical-documentation/complete-tech-docs/summary-file/sf2.pdf.

U.S. Census Bureau (2012b, December). American Indian and Alaska Native Summary File dataset. https://www.census.gov/data/datasets/2010/dec/aian-summary-file.html.

U.S. Census Bureau (2012c, May). National demographic analysis tables: 2010 (revised). https://www.census.gov/data/tables/2012/demo/popest/revised-2010-demographic-analysis-estimates.html; revised from original release in December 2010.

U.S. Census Bureau (2015, November 2015). *2020 Census Operational Plan: A New Design for the 21st Century. Version 1.1.* Washington, DC: U.S. Census Bureau. Version 1.1 is described as unchanged in content from Version 1.0, differing only in page and graphical format.

U.S. Census Bureau (2016a, September). *2020 Census Operational Plan: A New Design for the 21st Century. Version 2.0.* Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2016b, May 25). *Census Enterprise Data Collection and Processing (CEDCaP) COTS Capability Assessment and Analysis.* Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2017, September). *2020 Census Operational Plan: A New Design for the 21st Century. Version 3.0.* Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2018, December). *2020 Census Operational Plan: A New Design for the 21st Century. Version 4.0.* Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2019, October). 2010 demonstration P.L. 94-171 Redistricting Summary File and Demographic and Housing Demonstration File—technical documentation. Originally issued October 2019; user notes updated January 24, 2020. https://www2.census.gov/programs-surveys/decennial/2020/program-management/data-product-planning/2010-demonstration-data-products/01-Redistricting_File--PL_94-171/2019-10-29_Summary_File/00-README_2010_Demonstration_Data_Products/01-2010_Demonstration_Data_Product_Technical_Document.pdf.

U.S. Census Bureau (2021a, April 26). 2020 Census Operational Quality Metrics, Release 1. https://www2.census.gov/programs-surveys/decennial/2020/data/operational-quality-metrics/census-operational-quality-metrics-release_1.xlsx.

U.S. Census Bureau (2021b, May 28). 2020 Census Operational Quality Metrics, Release 2. https://www2.census.gov/programs-surveys/decennial/2020/data/operational-quality-metrics/census-operational-quality-metrics-release_2.xlsx.

U.S. Census Bureau (2021c, August 25). 2020 Census Operational Quality Metrics, Release 3, Table 2 (Item Nonresponse Rates). https://www2.census.gov/programs-surveys/decennial/2020/data/operational-quality-metrics/census-operational-quality-metrics-release_3-table2.xlsx.

Copyright National Academy of Sciences. All rights reserved.

U.S. Census Bureau (2021d, June). *2020 Census State Redistricting Data (Public Law 94-171) Summary File, 2020 Census of Population and Housing—Technical Documentation*. SFSRD/20-01. Washington, DC: U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/2020/technical-documentation/complete-tech-docs/summary-file/2020Census_PL94_171Redistricting_StatesTechDoc_English.pdf.

U.S. Census Bureau (2021e, November). *Disclosure Avoidance for the 2020 Census: An Introduction*. Washington, DC: U.S. Government Publishing Office.

U.S. Census Bureau (2021f, December). Methodology for the United States population estimates: Vintage 2021—nation, states, counties, and Puerto Rico—April 1, 2020 to July 1, 2021. https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/2020-2021/methods-statement-v2021.pdf.

U.S. Census Bureau (2021g, August 12). Supplementary tables on race and Hispanic origin: 2020 Census redistricting data (P.L. 94-171). https://www.census.gov/data/tables/2020/dec/2020-redistricting-supplementary-tables.html.

U.S. Census Bureau (2022a, August). 2020 census data product planning crosswalk. Excel spreadsheet. https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-development.html.

U.S. Census Bureau (2022b, January). *2020 Census Operational Plan: A New Design for the 21st Century. Version 5.0.* Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2022c, March 10). Census Bureau releases estimates of undercount and overcount in the 2020 Census. Press release CB22-CN.02. https://www.census.gov/newsroom/press-releases/2022/2020-census-estimates-of-undercount-and-overcount.html.

U.S. Census Bureau (2022d, March 10). National demographic analysis tables: 2020. https://www.census.gov/data/tables/2020/demo/popest/2020-demographic-analysis-tables.html.

U.S. Census Bureau (2023a, September 13). 2020 Census data product—Demographic and Housing Characteristics File A (Detailed DHC-A). Pre-release webinar, presentation slides. https://www.census.gov/content/dam/Census/newsroom/press-kits/2023/detailed-dhc-a/20230913-webinar-slides-detailed-dhc-a.pdf.

U.S. Census Bureau (2023b, May 31). Census Bureau provides update on 2020 Census data products. Press release CB23-CN.11. https://www.census.gov/newsroom/press-releases/2023/2020-census-data-products.html.

U.S. Census Bureau (2023c, April 12). Workshop on assessing fitness-for-use of differential privacy-adjusted census data. Presented at Population Association of America meetings, Atlanta, GA.

Copyright National Academy of Sciences. All rights reserved.

U.S. Census Bureau, Population Division (2012, July 5). Modified race summary file methodology statement. https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/modified-race-summary-file-method/mrsf2010.pdf.

U.S. Government Accountability Office (2020, February 12). *2020 Census: Inital Enumeration Underway but Readiness for Upcoming Operations is Mixed*. GAO-20-368R. Washington, DC: U.S. Government Accountability Office.

U.S. Government Accountability Office (2023, March). *2020 Census: A More Complete Lessons Learned Process for Cost and Schedule Would Help the Next Decennial*. GAO-23-105819. Washington, DC: U.S. Government Printing Office.

Van Riper, D., T. Kugler, and J. Schroeder (2023). IPUMS NHGIS privacy-protected 2010 Census demonstration data [databases]. https://www.nhgis.org/privacy-protected-2010-census-demonstration-data. Minneapolis, MN.

Velkoff, V. and E. Jensen (2022). *Evaluating the Demographic Quality of the United States 2020 Census*. Note for Meeting of Group of Experts on Population and Housing Censuses, September 21–23, 2022. Geneva: United Nations Economic Commission for Europe and Conference of European Statisticians. https://unece.org/sites/default/files/2022-09/WP7_USA_Velkoff_ENG.pdf.

Villa Ross, C. (2023, June). *Uses of Decennial Census Programs Data in Federal Funds Distribution: Fiscal Year 2021*. Washington, DC: U.S. Census Bureau. https://www2.census.gov/library/working-papers/2023/decennial/census-data-federal-funds-fy-2021.pdf.

Vink, J. (2023, February 9). Usability of age distributions in the DAS demonstration data. Presented at Applied Demography Conference, Annapolis, MD.

Wagner, D. and M. Layne (2014). The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) record linkage software. Center for Administrative Records Research and Applications Working Paper 2014-01, U.S. Census Bureau, Washington, DC.

Werner, A. (2022, June 21). Evaluating the potential impacts of differential privacy on public health—a series of use cases from the Centers for Disease Control and Prevention. Presentation at Workshop on the Demographic and Housing Characteristics Files, Washington, DC. https://www.nationalacademies.org/event/06-21-2022/2020-census-data-products-workshop-on-the-demographic-and-housing-characteristics-files.

West, K. K., J. G. Robinson, and M. Bentley (2005). Did proxy respondents cause age heaping in the census 2000? In *Proceedings of the Section on Survey Research Methods*, pp. 3658–3665. Alexandria, VA: American Statistical Association.

Whitford, D. C. (2008, July 9). *Overview of 2010 Census Coverage Masurement Program*. DSSD 2010 Census Coverage Measurement Memorandum Series #A-19. Washington, DC: U.S. Census Bureau. https://www2.census.gov/programs-surveys/decennial/2010/program-management/5-review/a-19.pdf.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Wright, T. and K. Irimata (2021, August 5). *Emirical Study of Two Aspects of the Topdown Algorithm Output for Redistricting: Reliability & Variability (August 5, 2021 Update)*. Study Series—Statistics #2021–02. Washington, DC: U.S. Census Bureau.

Yorsz, W. (1976, May). A study of the movement for a national statistical data center (1960–1967). Master's thesis, Massachusetts Institute of Technology. https://dspace.mit.edu/bitstream/handle/1721.1/60107/02360954-MIT.pdf.

Zelnik, M. (1961). Age heaping in the United States census: 1880–1950. *The Milbank Memorial Fund Quarterly 39*(3), 540–573. https://doi.org/10.2307/3348729.

Copyright National Academy of Sciences. All rights reserved.

Appendixes

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

– A –

Full List of Conclusions and Recommendations

CONCLUSIONS AND RECOMMENDATIONS FROM THE INTERIM REPORT

Conclusion INT-4.1: The 2020 Census was implemented in light of severe and unprecedented operational challenges. Faced with extremely difficult decisions and reconciling operational demands with strong and appropriate concern for public health conditions, the professional staff of the Census Bureau generally made principled and reasoned choices in adjusting 2020 Census operations to the COVID-19 pandemic, natural disasters, and other disruptions. The basic fact that the 2020 Census was completed, as close to schedule as it was, is itself a major accomplishment, and the Census Bureau and its staff (and the responding American public) deserve commendation for heroic effort amidst the difficult circumstances.

Conclusion INT-4.2: The ability of the Census Bureau to complete the 2020 Census amidst its difficult circumstances depended critically on early commitment in the preceding decade to a general design for the 2020 Census, premised on targeted development work in a tractable number of priority innovation areas: increased field automation, wider use of administrative records data in census processes, modernized address list development, and Internet response.

Copyright National Academy of Sciences. All rights reserved.

Conclusion INT-4.3: The fact that the 2020 Census was completed under difficult circumstances is not *the same, and is not meant to be interpreted, as a broader statement that the 2020 Census and its data products are high quality and credible. While that may be the case, the evidence base for such a definitive statement on 2020 Census quality is still in development, and such statements require careful attention to bases of comparison.*

Conclusion INT-4.4: The 2020 Census labored under formidable challenges, disruptions, and impediments, many of which were forced upon the Census Bureau in the sense that they were purely beyond the Bureau's control. But it must also be acknowledged that some of the disruptions and impediments to the 2020 Census arose from design decisions and factors within the Census Bureau's control.

Conclusion INT-4.5: The 2020 Census partnership and communication programs and the mobilization of complete count committees and other outreach/get-out-the-count efforts (including those with states and philanthropic bodies) were intended to boost public attention to the census, but there has been little attention to the return on investment and effectiveness of these support operations.

Conclusion INT-4.6: It will not be possible for this panel (or any other evaluator) to understand and characterize the quality of the 2020 Census unless the Census Bureau is forthcoming with informative data quality metrics, including new measures based on operational/process paradata, at substate levels and small-domain spatiotemporal resolution, unperturbed by noise infusion.

Recommendation INT-4.1: The Census Bureau should work on ways to make 2020 Census data quality metrics publicly available at small-domain spatiotemporal resolutions, unperturbed by disclosure avoidance, to bolster confidence in the published tabulations. The Census Bureau should also develop ways to enable qualified researchers to access a full range of data quality metrics and report their findings.

Recommendation INT-4.2: The Census Bureau should cast its operational assessments and evaluations of the 2020 Census as evidence of the effectiveness of census operations and key inputs to the research agenda for the 2030 Census, rather than purely procedural and documentary of the 2020 experience. Historically, the Census Bureau's operational assessments and evaluations have been limited to high-level tabulations; the 2020 program should examine 2020 Census operations at finer levels of spatiotemporal and demographic resolution.

Copyright National Academy of Sciences. All rights reserved.

OVERVIEW OF THE 2020 CENSUS

Recommendation 2.1: In addition to completing its own program of evaluation and assessment reports, the U.S. Census Bureau should complete and publish a comprehensive procedural history of the 2020 Census.

AGE HEAPING AS AN INDICATOR OF DATA QUALITY

Conclusion 3.1: Age heaping (excess responses of ages ending in 0 or 5) is a well-recognized indicator of data quality in censuses. Age heaping was much more prevalent in the 2020 Census than in 2010, by an estimated factor of at least 2.5 for ages between 23 and 62 years, for which heaping is most common. This increase was almost completely due to increases in the rate of heaping during Nonresponse Followup (NRFU)—especially from proxy respondents but also from NRFU household respondents. Age heaping was rare in both censuses for self-responses, whether by paper or internet (in 2020 only), and there was no evidence of age heaping associated with administrative records.

Conclusion 3.2: Findings about age heaping raise concerns about the overall quality of data collected during 2020 Nonresponse Followup (NRFU), especially from proxy respondents but even from some household members. Other quality indicators, including item nonresponse rates, numbers of people requiring imputation of all characteristics, and coverage errors, reinforce the difficulties of obtaining accurate data in NRFU and indicate that the difficulties increased in 2020 compared with 2010. The greater extent of quality problems in NRFU in 2020 compared with 2010 may be attributable to the COVID-19 pandemic (including the resulting delay in the NRFU schedule), but other factors might also be involved, which would need to be addressed in 2030 Census planning.

COVERAGE MEASUREMENT IN THE 2020 CENSUS

Conclusion 4.1: Comparing the 2010 and 2020 Census net population coverage estimates from Demographic Analysis (DA):

- *For total population, both censuses, based on the middle series of DA estimates, closely approximated the estimated population total (census erroneous enumerations almost offset census omissions).*
- *For age and sex, the 2020 Census undercounted children ages 0–4 more than the 2010 Census. The 2020 Census also overcounted*

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

college-age people substantially more than the 2010 Census and, similar to the 2010 Census, overcounted people ages 50 and older.

- *To date, no estimates of net coverage error for the 2020 Census by race are available because production of the necessary census race data was delayed by development of confidentiality-protection routines.*

Conclusion 4.2: Comparing the 2010 and 2020 Census population coverage estimates from the Post-Enumeration Survey (PES):

- *Differences among race and ethnic groups widened substantially in 2020, with adverse implications for use of the data to allocate fixed resources, such as representation, funding, and services. Some groups saw increases in net overcounts, particularly non-Hispanic White Alone people and Asian people, and other groups saw increases in net undercounts, particularly Black people, Hispanic people, and American Indians and Alaska Natives on reservations.*
- *Increases in net undercounts in 2020 for groups such as Black and Hispanic people were driven by increases in omissions from the census (as measured by a lower match rate for people in the P-sample to the E-sample in the PES). Omissions, in turn, were higher in census tracts with lower rates of self-response.*
- *Net overcounts and undercounts were more pronounced for people of various groups in rental compared with owner housing, particularly in 2020 compared with 2010.*

Conclusion 4.3: With respect to the quality of the 2020 Post-Enumeration Survey (PES) and Demographic Analysis (DA) estimates:

- *The quality of the 2020 PES deteriorated compared with 2010, due to smaller sample sizes, higher nonresponse rates, higher missing data rates, higher rates of unresolved match status and unresolved enumeration status, and delays caused by the COVID-19 pandemic.*
- *While DA continued to set a standard for estimation of the total population in 2020, it showed a wide range of net coverage estimates, based on Medicare records and assumptions about completeness of coverage, for people ages 75 and older among the low, middle, and high series of estimates. The 2020 DA coverage estimates series for Hispanic and non-Hispanic children and young adults also showed a wide range, largely due to differences in metnods for assigning Hispanic births. Differences in net international migration assumptions also likely played a role, particularly for older ages.*

Copyright National Academy of Sciences. All rights reserved.

Recommendation 4.1: For the Demographic Analysis (DA) coverage evaluation method, the U.S. Census Bureau should conduct research on:

- **The reasonableness of the assumption for incompleteness of Medicare enrollment in the high series, which produced coverage estimates that diverged greatly from the low and middle series for people ages 75 and older, through an appropriate match study;**
- **The suitability of Medicare data for coverage evaluation of the entire population ages 65 and older as in the 2010 Census and for improving the census count of this age group in 2030;**
- **The reasonableness of the methods for assigning Hispanic and non-Hispanic births for coverage estimates for children and young adults, which produced wide differences between the low and high DA series in 2020;**
- **Methods for narrowing estimates of net international migration, which affect estimates for Hispanic people and the total population; and**
- **Methods for developing experimental subnational DA estimates, starting with young children.**

Recommendation 4.2: The U.S. Census Bureau should conduct additional analyses of the 2020 Post-Enumeration Survey (PES) to learn as much as possible about census omissions. One such analysis should include P-sample and E-sample households with at least one member in common, to assess not only erroneous census enumerations but also omissions and how they vary with household characteristics and type of enumeration (e.g., proxy interview, internet response). In addition, the Census Bureau should perform discriminant or similar analyses to identify variables that contributed to census component errors (e.g., omission, duplication) for 2020 P-sample and E-sample addresses, households, and individuals, which could identify census processes to target for improvement in 2030.

Recommendation 4.3: For the Post-Enumeration Survey (PES) method, the U.S. Census Bureau should:

- **Perform sensitivity analyses, based on plausible assumptions, to put error bounds around such key operations as matching and imputation of match and enumeration status, to evaluate the quality of the 2020 PES and plan such analyses from the outset for the 2030 PES;**

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

- **Plan analyses for the 2030 PES of components of error for census operations in exhaustive and mutually exclusive categories that are as comparable as possible between 2020 and 2030;**
- **Seek adequate funding to increase the PES sample size in 2030 to at least 2010 levels; and**
- **Experiment with modeling to estimate net undercoverage and overcoverage for more detailed geographic and demographic groups for which direct estimates could not be provided in 2020.**

Recommendation 4.4: The U.S. Census Bureau should prioritize research on potential sources of coverage errors—both undercounts and overcounts—for geographic areas and population groups, using additional methods besides further analysis of the Post-Enumeration Survey. To address omissions from the census, the Census Bureau should match 2020 Census records with the 2019–2021 American Community Survey and a wide variety of administrative records for census tracts—perhaps sampling those with low self-response rates. The goal would be to provide an evidence base for methods and data sources that could potentially reduce omissions in the 2030 Census. To address duplications in the census, the Census Bureau should match 2020 Census records with themselves to identify duplicate enumerations of people with more than one residence, college students enumerated at college and at home, children in joint custody arrangements, and the like. The goal would be to provide an evidence base for methods to reduce duplicates in the 2030 Census and potentially to remove them from the count, just as people duplicated within a household are dropped from the count.

MASTER ADDRESS FILE

Conclusion 5.1: The U.S. Census Bureau's approach of retaining a fluid Master Address File database and filtering operational extracts from that database is a sound one, provided that the filters are continually studied and updated to improve their efficacy.

Conclusion 5.2: The U.S. Census Bureau's reliance on the U.S. Postal Service's Delivery Sequence File as a primary update source between censuses is a good one, though it necessitates a reliable update source in areas without mail delivery.

Copyright National Academy of Sciences. All rights reserved.

Conclusion 5.3: With respect to the user-input address-building operations of Local Update of Census Addresses, Count Review, and New Construction, both participant governments and the U.S. Census Bureau erred on the side of inclusion—with local authorities submitting many more addresses for review than were actually in the operation's scope and the Census Bureau injecting many of those submitted addresses into the enumeration process to let fieldwork determine their accuracy. Consequently, these user-input operations appear to have added more bad (unenumerated) addresses to the 2020 Census than good (enumerated) addresses.

Conclusion 5.4: The Office-Based Address Verification operation for resolving respondent-supplied addresses from Non-ID Processing was an effective innovation in the 2020 Census, both in terms of its potential for resolving address listing issues with limited field work and for informing revisions to Master Address File filters.

Recommendation 5.1: The U.S. Census Bureau should continue to research and refine the filters it applies to the Master Address File to derive functional operational extracts, with the intent to reduce the number of addresses cancelled during collection (i.e., flagged as deletes in Nonresponse Followup and other field operations). Along those lines, the Census Bureau should continue to study ways to partner with and more fully utilize U.S. Postal Service data (including the Delivery Sequence File) and potential sources for address addition and revision information between censuses, such as address data recoverable from administrative records extracts, with particular eye toward more regular updating in areas without mail delivery.

Recommendation 5.2: The rules and scope of participant-input address-building operations of the census require clarification, including the Local Update of Census Addresses, New Construction, and Count Review. The U.S. Census Bureau should continue to improve ways to work with user-supplied input or data resources, and make it easier for state, local, and tribal authorities to supply input in usable form, while more clearly laying out the expectations for those address-building operations.

Recommendation 5.3: The U.S. Census Bureau should extend and improve during-the-decade programs, akin to the Geographic Support Systems Initiative of the 2010s, with the goal of continuous address- and map-database refinement. Such

Copyright National Academy of Sciences. All rights reserved.

efforts should include engagement with state, local, and tribal partners, including on the question of whether some addresses can safely be removed from the Master Address File.

Recommendation 5.4: Lessons learned from the Office-Based Address Verification operation should be documented and incorporated into the tools and techniques for ongoing, continuous address and map data refinement (e.g., continuous In-Office Address Canvassing) as well as revisions of the Non-ID Processing operation.

Recommendation 5.5: The U.S. Census Bureau should research and test approaches to increase the proportion of dwellings that can be mailed to while still being associated with a physical geographical location. This would reduce the proportion of addresses in Update Leave and expand the use of approaches yielding higher Self-Response for more dwellings in the 2030 Census.

SELF-RESPONSE TO THE CENSUS

Conclusion 6.1: Going from a paper-only 2010 Census to a majority Internet Self-Response 2020 Census is a commendable achievement, and it is particularly commendable that the U.S. Census Bureau's internally developed Primus system was able to be swapped in immediately prior to the census with excellent performance throughout the count. (Primus was the backup system pressed into primary service after concerns surfaced regarding the outsource-developed system.) The distinction between Internet First and Internet Choice contact strategies was a reasonable starting position.

Conclusion 6.2: With available return rate and operational data, it was impossible to fully evaluate the 2020 Census mailing and contact strategy beyond gross effects, such as somewhat earlier elevated levels of paper-questionnaire takeup in Internet Choice areas, and it was also impossible to disentangle effects of messaging, delivery, and response mode choices brought about by operations during the COVID-19 pandemic.

Conclusion 6.3: Mode configurations varied substantially across geography, and this variation was associated with social, economic, and housing characteristics. People in areas with higher concentrations of minority populations, lower socioeconomic status, and less-stable housing were more likely to be enumerated in Nonresponse Followup and by modalities with high item nonresponse. The quality of 2020

Copyright National Academy of Sciences. All rights reserved.

Census data (as indicated by higher item nonresponse) is poorer in areas with higher concentrations of minorities, lower socioeconomic status, and less stable housing. This raises concerns about data equity, with progressively less accurate and less complete data being collected differentially by race and ethnicity, socioeconomic status, and housing tenure.

Conclusion 6.4: The magnitude of uncertainty in terms of the accuracy of counts of population characteristics as a function of item nonresponse is unknown. The variation in enumeration modality associated with social, economic, and housing characteristics results in differences in data quality by these characteristics. We do not understand the overall impact of item nonresponse and the associated imputation of values on the accuracy of data for the total population or subpopulations. A better understanding of the differential impact of imputation across subpopulations would potentially enable prioritized targeting of groups, to improve self-reporting.

Recommendation 6.1: The U.S. Census Bureau should continue to investigate the performance of the contact strategies used in the 2020 Census, to increase overall self-response (particularly via the internet) in 2030, more in concordance with international experience with online census response. This investigation should include review of the effectiveness of the 2020 Census's segmentation into Internet First and Internet Choice strategies, as well as the timing and the operative contact/enumeration strategy (e.g., Internet First or Internet Choice mailings or Update Leave delivery) underlying successful Non-ID enumerations in 2020.

Recommendation 6.2: The U.S. Census Bureau should engage in further research to more specifically identify social, economic, and housing characteristics associated with less-than-accurate forms of enumeration, and should also research related communication and operational strategies that would improve self-response and associated data accuracy for those population and housing segments associated with poorer-quality census data in 2020.

Recommendation 6.3: The U.S. Census Bureau should assess the magnitude of the impact of imputation on the accuracy of 2020 Census data. Assessment of the differential impact of imputation on subpopulations should be used to prioritize research into strategies to improve self-response.

Copyright National Academy of Sciences. All rights reserved.

NONRESPONSE FOLLOWUP

Conclusion 7.1: Analysis of the resolution of Nonresponse Followup (NRFU) cases by operational phase does not, in itself, show signs that 2020 Census data collection was unduly hastened to completion. Enumeration by proxy report or by administrative records (if records information of sufficient quality was available) was invoked in all stages of NRFU, but not in a way suggestive of an egregious rush to closeout.

Conclusion 7.2: Though the logic of the U.S. Census Bureau's split of Nonresponse Followup (NRFU) into distinct operational phases is generally sound, it is impossible to assess the effectiveness of the phased approach as originally designed because the circumstances of the 2020 Census precluded the planned winnowing of NRFU staff. The determination of change from one NRFU phase to another was data driven, based on resolution of a specified percentage of cases in a particular geographic area, which is appropriate. However, the specific operational geography at which NRFU phases were determined—Census Field Supervisor Areas—was so operational and tactical in nature, and so divorced from familiar tabulation geography, that it is unworkable from an evaluation standpoint and raises concerns that it might be difficult for census managers to appreciate actual completion coverage during live field operations.

Conclusion 7.3: The information available for the panel's analysis did not permit assessment of the potential effects of the delays in 2020 Nonresponse Followup (NRFU) and other field operations in 2020 on the quality of results, such as recall biases and the coding of accurate occupancy statuses of dwellings (especially later in the NRFU period).

Conclusion 7.4: Nonresponse Followup (NRFU) enumerations result in more item nonresponse and poorer coverage than self-responses. NRFU item nonresponse for race and Hispanic origin exceeded 15% in the 2020 Census. Proxy enumerations had greater than 35% item nonresponse for race and Hispanic origin. Results suggest that census tracts with higher proportions of people with lower socioeconomic status (SES) characteristics were more likely to be enumerated using modes that resulted in higher item nonresponse. The net result is that 2020 Census data on race, Hispanic origin, and age are less accurate in those census tracts than in those with higher proportions of higher SES characteristics. This result is most relevant for redistricting. State and local officials are defining political districts with data that vary in how accurately they reflect the age and racial and ethnic composition of the population. Census data are particularly prone to error in areas

Copyright National Academy of Sciences. All rights reserved.

with high percentages of people who are Black, Hispanic, or American Indian and Alaska Native living on reservations.

Recommendation 7.1: As a goal for the 2030 Census, the U.S. Census Bureau should consider a major reduction in the use of proxy interviewing for enumeration, if not the elimination of proxy reporting in all but very limited circumstances. Work toward this goal should be predicated on the results of research that only the Census Bureau can conduct, reexamining and comparing proxy reports in the 2020 Census with the information that would have been available from administrative records or third-party data resources.

Recommendation 7.2: The U.S. Census Bureau should consider conducting research on the quality, and factors associated with variation in the quality, of proxy and count-only responses, with the aim of minimizing the occurrence of these factors.

Recommendation 7.3: The U.S. Census Bureau should conduct fine-grained analysis of Nonresponse Followup enumerations (separately for household, proxy, and administrative records responses) for small geographic areas and population groups as part of its research to develop and test ways to improve the quality of enumeration in the 2030 Census. The Census Bureau should publish the results of its analysis as an aid to users of 2020 Census data.

USE OF ADMINISTRATIVE RECORDS FOR ENUMERATION IN THE 2020 CENSUS

Conclusion 8.1: Based on the information currently available, the U.S. Census Bureau deserves commendation for the meaningful steps made in the limited use of administrative records as an alternative to the enumeration of some nonresponding households in the 2020 Census. The Census Bureau acted with appropriate levels of both boldness and caution in settling on conditions under which a nonresponding household at a designated residence would be enumerated using records information after only one enumerator visit rather than six, yet still ensuring that single in-person visit. Moreover, this change in approach was preceded by an extensive history of high-quality experimentation and testing.

Conclusion 8.2: Moving to the use of administrative records to enumerate a substantial proportion of the population without a contact attempt does not seem feasible for the 2030 Census. Rather, it appears

Copyright National Academy of Sciences. All rights reserved.

to be a long-term proposition, requiring extensive testing, research, and engagement with Congress, other stakeholders, and the public. Not only are there operational and technical issues to resolve and data-quality standards to determine and assess for geographic areas and population subgroups, but there is also the need to consider the ramifications of switching from a household/address-based focus to a person-based focus and to address issues of legal, constitutional, and public acceptability. It will be essential for the U.S. Census Bureau to be as transparent as possible about its research and testing and to have open and constructive dialogue with all parties.

Recommendation 8.1: The U.S. Census Bureau should continue its research and development program concerning the best ways to use administrative records as a supplement to decennial census operations. Potential uses of administrative records include expanding enumeration of limited subsets of the 2030 Census population in the Nonresponse Followup workload, reducing proxy responses and whole-person imputations, and possibly redressing the long-standing net undercount of children ages 0–4.

MEASUREMENT OF THE GROUP QUARTERS POPULATION

Conclusion 9.1: The enumeration of nonhousehold group quarters (GQ) locations in the 2020 Census experienced major shortcomings, some of which can be directly attributed to COVID-19 pandemic adjustments to the data-collection scheduling (e.g., the scuttling of prearranged plans for in-person enumeration worked out in the Group Quarters Advance Visit operation). Yet several important operational and conceptual shortcomings cannot be attributed solely to the pandemic, such as the mismatch between census content and the U.S. Department of Education's interpretation of permissible data provision under federal privacy law, the shortfalls in the enumeration of people in military GQs, as well as the fluctuation in status (and resulting problems in the address list) in more fluid GQ types like group homes.

Recommendation 9.1: The U.S. Census Bureau's steps to make the Master Address File an ongoing, continuous inventory of all living quarters—including group quarters (GQs)—should continue, with an eye toward developing updated information for GQs on par with the regular updating of conventional housing units. This should include a separate inventory of GQ types that have longstanding address provenance and those that

Copyright National Academy of Sciences. All rights reserved.

need to be tracked more closely because they are continually changing.

Recommendation 9.2: Group quarters (GQs) require more than updated address information; they require continually updated contact strategies as well. The U.S. Census Bureau should develop "selected advance contact" protocols with GQ facilities by type (and their parent companies, as appropriate), akin to follow-ups performed with businesses, and state, local, and tribal partners should be engaged in this update activity.

Recommendation 9.3: In light of the 2020 experience, the U.S. Census Bureau should convene group quarters (GQ) stakeholders in a discussion of more effective and useful options for electronic provision of data for GQ residents, and on the issues involved in obtaining access to administrative records for GQ facilities. The completeness of such records should also be assessed by GQ type.

MEASUREMENT OF RACE AND ETHNICITY

Conclusion 10.1: The 2020 Census depicted a more diverse nation than did the 2010 Census. The White Alone population declined significantly, particularly among Hispanic people; the Hispanic population increased; and the Two or More Races population increased substantially, particularly among Hispanic people. It is unknown the extent to which these changes were due to demographic changes (e.g., more multiracial children); changes in self-identification; the provision of write-in origins (perhaps not intended to indicate racial identification) made possible by the format and processing changes to the 2020 Census race question; and other effects of the 2020 Census format and processing changes.

Conclusion 10.2: The 2020 Census had poorer quality data on race and ethnicity compared with the 2010 Census in terms of coverage error (net overcounts and undercounts), rates of missing and imputed responses, and the noise infused by the 2020 Census Disclosure Avoidance System.

- *The 2020 Census exhibited larger coverage errors compared with 2010 for race and ethnic groups. There were increases in net overcounts for some groups, particularly non-Hispanic White Alone people and Asian people, and increases in net undercounts for other groups, particularly Black people, Hispanic people, and American Indians and Alaska Natives. Consequently, differences among groups widened significantly in 2020, with*

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

adverse implications for uses of census data to allocate fixed resources, such as representation, funding, and services.

- *The 2020 Census had higher imputation rates (to account for missing and inconsistent responses) for race and ethnicity compared with 2010, particularly for responses not provided by a household member via the internet, a mailed-back paper questionnaire, or telephone. Moreover, the 2020 Census had much higher imputation rates for race and ethnicity for people in group quarters compared with the already high rates in 2010.*
- *The 2020 Disclosure Avoidance System injected noise of considerable magnitude into the Redistricting File for small race and ethnicity groups in small governmental jurisdictions, such as incorporated places and American Indian tribal lands and Alaska Native villages.*

Recommendation 10.1: The U.S. Census Bureau should conduct research to determine how the changes in format and processes that were made in the 2020 Census and in the American Community Survey beginning in 2020 affected the distributions of race and ethnicity. Such research should use qualitative, quantitative, and simulation methods to ascertain: how respondents viewed and used the 2010 and 2020 formats; trends in multirace reporting by age, sex, race, and ethnicity; how samples of 2020 respondents would have been categorized using the 2010 format, data capture, and coding; and the implications of differences in write-ins by response mode (e.g., more write-ins for internet responses) on race distributions among geographic areas. The Census Bureau should communicate the results of this research to data users to assist them to understand the implications of the changes for time series, and should also use the findings to inform 2030 Census planning.

Recommendation 10.2: To improve the quality of reporting on race and ethnicity in the census, American Community Survey (ACS), and other federal data collections, the U.S. Office of Management and Budget (OMB) should revise Statistical Policy Directive No. 15 to adopt a combined, check-more-than-one race/ethnicity classification with both Hispanic and Middle Eastern or North African as main categories, in addition to White, Black, American Indian and Alaska Native, Asian, and Native Hawaiian and Other Pacific Islander. OMB should formalize any changes to the race and ethnicity reporting standards as soon as practicable, to permit the U.S. Census Bureau to implement the new race/ethnicity question in the

Copyright National Academy of Sciences. All rights reserved.

ACS and in the 2030 Census. OMB should allow agencies to collect data only for the major categories in the combined question when added detail (as in the 2020 Census) would impose undue administrative burdens. A decision by the Census Bureau on including added detail and write-in spaces, along with expanded data capture and coding procedures, for a combined race/ethnicity question for the ACS and the 2030 Census should await completion of research on such matters as how respondents viewed the 2020 question format and the effects of response-mode differences in write-in responses on diversity among geographic areas.

Recommendation 10.3: The U.S. Census Bureau should produce a crosswalk or bridge between the 2010 and 2020 Census race and ethnicity questions and responses. Similarly, if a combined race/ethnicity question (with Hispanic and Middle Eastern or North African categories) is adopted as the standard by the U.S. Office of Management and Budget, the Census Bureau should produce a crosswalk or bridge between the 2020 version and revised race and ethnicity questions and responses, as soon as the revisions are implemented in the American Community Survey and then in the 2030 Census. The Census Bureau should involve data users in this important work.

Recommendation 10.4: The U.S. Census Bureau should consult as early as possible with the redistricting community, which includes state legislators, redistricting commissions, political parties, and political consultants, among others, to determine the optimum set of tabulations to include on the 2030 Redistricting File, whether the current race and ethnicity categories are retained or the proposed revisions take effect. The consultation should include consideration of streamlining the number of race/ethnicity categories in the file and combining blocks with small populations (using input from localities) to maximize the ability to protect confidentiality while maintaining accuracy.

IMPACT OF DISCLOSURE AVOIDANCE CHANGES ON 2020 CENSUS DATA PRODUCTS

Conclusion 11.1: In an era of Big Data, linkage technology, and other aspects of today's computational and data environment, it is difficult for the decennial census to balance the need for accurate data for small areas and small population groups with an adequate

Copyright National Academy of Sciences. All rights reserved.

level of confidentiality protection. When evaluating methods and levels of confidentiality protection, stewardship of census data for the public good requires balancing societal needs for the information—particularly for equity in representation, fund allocation, and many other uses—with plausible harms of breaches of confidentiality.

Conclusion 11.2: The decision of the U.S. Census Bureau to respond to the threats to confidentiality protection at a very late date in 2020 Census planning with a new, more complex Disclosure Avoidance System (DAS) using differential privacy-based algorithms went counter to long-standing principles of decennial census planning. The approach had not been tested in a census environment nor had the ability of the algorithms to handle critical user data needs been assessed; the Census Bureau had no backup plan should implementation of the new DAS prove challenging. The decision to continue to deploy the new DAS in the face of serious implementation problems has resulted in marked delays in delivery of data products, with some variables and types of geographic units of questionable utility, and other variables and geographies not provided at all. In addition, it is not clear that the chosen privacy budgets for the various 2020 Census data products, with high values of the ϵ parameter that trade off accuracy with confidentiality protection, provide much actual protection.

Conclusion 11.3: Sustained, interactive communication with data users in advance of major changes to data products from a census or long-running survey is critical for a statistical agency to maintain credibility with and serve its user community. The late decision to adopt new confidentiality-protection techniques for the 2020 Census led to untimely and inadequate user communication on the part of the U.S. Census Bureau. The comment periods after the release of each demonstration file for the Redistricting and Demographic and Housing Characteristics Files were valuable, but many users lacked the staff and resources to respond adequately.

Conclusion 11.4: By not releasing unperturbed operational quality metrics for census tracts, the U.S. Census Bureau missed an opportunity to inform the public of quality differences in the 2020 Census among substate areas and to gain the assistance of interested stakeholders in planning for 2030. Other than the theoretical concept that the release of any piece of information about an individual increases the risk of disclosure, the Census Bureau has not justified its decision to release census tract metrics that are too noisy for meaningful analysis.

Conclusion 11.5: For a major change in data products such as the new confidentiality-protection techniques for the 2020 Census, it is

Copyright National Academy of Sciences. All rights reserved.

essential for a statistical agency to document fully and publicly the need for the change and the implications for the products. In the case of the 2020 Census, there has not yet been peer review of the crucial 2016–2018 reconstruction-reidentification study that was the public basis for the decision to adopt a formal privacy approach for protecting confidentiality at a late date in 2020 Census planning. Similarly, there has not been an actual demonstration of why more-than-minimal confidentiality protection is needed for substate operational quality metrics (e.g., self-return rates for census tracts and counties) or why it is not possible to provide as-collected population counts to local governments with a reasonable level of confidentiality protection for other data.

Recommendation 11.1: For 2030 Census data products, the U.S. Census Bureau should adopt the risk-utility framework recommended by the Advisory Commission on Data for Evidence Building, which accepts that disclosure risk is a continuum, that not all data items are equally sensitive, and that federal statistical data need to be accessible and useful for a wide range of users and uses.

Recommendation 11.2: At a minimum, the 2030 Census should provide as-collected (i.e., unaltered) total population counts for all governmental units (states, counties and equivalents, minor civil divisions, incorporated places, recognized tribal areas, school districts) and quasi-governmental units (census county divisions, census-designated places, and tribal statistical areas), no matter how small. In addition, census block population totals should add up to block groups, census tracts, counties, and states.

Recommendation 11.3: For the 2030 Census data product plan, the U.S. Census Bureau should begin immediately on a multipronged research program with ample testing and opportunities for feedback and dialogue with the data user and stakeholder community, broadly defined. The goal should be an end-to-end plan by 2027–2028 for producing a suite of 2030 data products that serve user needs, are appropriately protected, and meet the time schedule of the 1990–2010 Censuses.

Recommendation 11.4: U.S. Census Bureau research and dialogue with the user community, broadly defined, on the end-to-end plan for the 2030 Census suite of data products should include:

Copyright National Academy of Sciences. All rights reserved.

- A review of the 2020 Census experience, including rigorous assessment of the Census Bureau's reconstruction-reidentification studies under plausible scenarios of potential attack and comprehensive debriefing of users.
- Consultation with the redistricting community on race/ethnicity and geographic detail (accounting for the possibility of a combined race/ethnicity question) that best trade off accuracy and confidentiality protection for the 2030 Redistricting File.
- Issuance of challenges to the research community, in addition to focused contracts, for research and development on a range of confidentiality-protection methods and understandable metrics of risk and accuracy to accompany those methods.
- Research on practical methods for users to account for noise injected into 2030 Census data by the selected confidentiality-protection techniques.
- Research on the application of cost-benefit and risk-utility analysis for making tradeoffs between confidentiality protection and utility of census data products.
- Research on the sensitivity of individual data items to relevant communities at differing levels of geographic aggregation and the implications for confidentiality protection.

Recommendation 11.5: The U.S. Census Bureau should welcome initiatives to add language to appropriate legislative vehicles, such as the Foundations for Evidence-Based Policymaking Act of 2018, that prescribes responsibilities and penalties for data users, in addition to agency staff, for willful, harmful disclosure of confidential information.

LEARNING FROM 2020, PREPARING FOR 2030

Recommendation 12.1: To prioritize its testing and development of both procedures and technical systems for the 2030 Census, the U.S. Census Bureau should take the following as primary objectives:

- Maximize self-response to the census, including better matching of contact and communication strategies to the desired response mode, with particular attention to hard-to-reach, at-risk populations;

Copyright National Academy of Sciences. All rights reserved.

- **Improve the quality of data in Nonresponse Followup, including reducing if not eliminating the use of low-quality proxy reporting when there is an alternative available;**
- **Reduce gaps in coverage and data quality associated with race, ethnicity, and socioeconomic status;**
- **Improve the quality of address listings and contact strategies for all living quarters, including group quarters; and**
- **Realign the balance between utility, timeliness, and confidentiality protection in 2030 Census data products.**

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

– B –

Glossary and Abbreviations

2010 Demonstration Data Products (DDP): A series of data files released by the Census Bureau, beginning in late 2019, that replicate some of the content of previously published 2010 Census data tables but that are generated by applying the 2020 Census *Disclosure Avoidance System* to the original 2010 Census data.

ACDEB: Advisory Committee on Data for Evidence Building

ACO: *Area Census Office*

ACS: *American Community Survey*

Active Block Resolution: In the original plan for *In-Office Address Canvassing* for the 2020 Census, a program that permitted Census Bureau analysts to virtually canvass "active blocks" flagged by *Interactive Review* through available information resources, such as state and local government property databases and commercial street-view imagery. In-office analysts could document and correct housing unit coverage problems based on the available evidence, permitting those blocks to be reclassified as "passive" and kept out of the *In-Field Address Canvassing* workload. Active Block Resolution began at the Census Bureau's *National Processing Center* in spring 2016 but was discontinued in early 2017 due to budgetary constraints.

AdCan: *Address Canvassing*

AdRec; AdRecs: *administrative records*

Address Canvassing (AdCan): A major census operation to improve the completeness of the *Master Address File*, conducted in the year before the 2000, 2010, and 2020 Censuses. The operation is particularly focused on

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

addresses that will be included in the main census mailings, that is, census blocks in the Mailout/Mailback (2000 and 2010) and Self-Response (2020) *type of enumeration area*. The 2000 Census Block Canvassing operation and the 2010 Census Address Canvassing operation were both intended to send field staff to every address and street in every block in the designated areas, with the 2010 Address Canvassing operation being the only 2010 operation to utilize handheld computers for data collection (thus acquiring map-spot locations for addresses whenever possible). The 2020 Census Address Canvassing operation had both In-Office and In-Field components, with the In-Office program further divided into several subcomponents (most prominently, *Interactive Review* and *Active Block Resolution*) until some of the subcomponents were discontinued for budgetary reasons. In support of the 2020 Census, In-Field Address Canvassing operated between August–October 2019, with temporary-hire listers carrying out the field operation using laptop computers (and not the smartphones used in later 2020 Census field operations).

Administrative records (AdRec; AdRecs): Records collected as part of the operation of federal, state, and local programs, typically fund allocation and tax programs. A prominent part of the research program of the 2010 and 2000 Censuses, administrative records were given wider use in the 2020 Census, both for the identification of probable vacant housing units and for the enumeration of occupied housing units (in lieu of multiple field visits in some cases and as a last resort in others).

Age heaping: Phenomenon in censuses and surveys in which respondent age is reported at certain levels at the expense of other levels; in U.S. experience, age heaping typically refers to excess reporting of round-sounding ages ending 0 or 5.

AIAN: American Indian or Alaska Native

American Community Survey (ACS): The U.S. Census Bureau's continuous survey program to measure detailed socioeconomic and other data that now operates in parallel with the decennial census. In the 2010 Census, the ACS formally replaced the long-form sample of households that had been administered a more detailed questionnaire in previous decades, making the 2010 and 2020 Censuses short-form only.

Apportionment: The reallocation of the 435 seats in the U.S. House of Representatives, based on population size, using data from a new decennial census. Under current law, apportionment is conducted using the method of *equal proportions*.

Area Census Office (ACO): In the 2020 Census, a temporary office established to support decennial census field operations, in turn overseen by one of the temporary Regional Census Centers; in the 2010 Census, these offices were known as Local Census Offices. The 2020 Census used a field infrastructure consisting of 248 Area Census Offices divided between 6 Regional Census

Copyright National Academy of Sciences. All rights reserved.

Centers, generally co-located in cities containing permanent Census Bureau Regional Offices (Atlanta, Chicago, Los Angeles, New York, and Philadelphia, with the Denver region's Regional Census Center being established in Dallas). By comparison, the 2010 Census used 494 Local Census Offices across 12 Regional Census Centers (the Census Bureau consolidated to six operational regions in 2012).

ASA: American Statistical Association

Basic Collection Unit (BCU): The smallest unit of geography used in 2020 Census data collection. Previous censuses designated separate "collection blocks" and "tabulation blocks," the former being used internally during data collection and the latter being the final, published identifiers for census blocks. The BCU concept largely inherited its form and labels from the vintage of census blocks (or block clusters) then in use by the Census Bureau for census testing and other survey programs. Importantly, *type of enumeration area* is assigned to each BCU.

BCU: *basic collection unit*

CBAMS: Census Barriers, Attitudes, and Motivators Survey; family of Census Bureau experiments in recent decennial censuses to elicit information on factors driving response and nonresponse to the census.

CCF: *Coverage Correction Factor*

CEDCaP: Census Enterprise Data Collection and Processing

CEF: *Census Edited File*

Census Edited File (CEF): The third major step in the product flow for post data collection processing in the 2020 Census, following the *Census Unedited File*. The CEF includes additional edits and imputation (affecting the characteristics of persons or housing units but not the overall counts), and is effectively the complete data set for all items in the census. The CEF is processed by the *Disclosure Avoidance System* to protect confidentiality of census responses; in the 2020 Census, it is the resulting Microdata Detail File that is used for tabulation of census products, rather than the CEF.

Census Field Supervisor Area (CFSA): Operational level of geography developed to distribute field staff work. In the 2020 Census *Nonresponse Followup* operation, changes in phase between full automation, semipermanent assignment, and closeout (phases 1–3 in the new structure for the program) were determined by case completion rates at the CFSA level. As a purely operational, not cartographic, geography, CFSAs do not correspond neatly with other similar geographies such as tracts.

Census Program of Evaluations and Experiments (CPEX): In recent U.S. censuses, the slate of formal operational assessments, evaluation reports, and major experiments conducted in conjunction with the census. The set of census coverage measurement activities (*Demographic Analysis* and the *Post-Enumeration Survey*)—and their results—are not themselves generally

Copyright National Academy of Sciences. All rights reserved.

under the CPEX umbrella, but the operational conduct of specific Post-Enumeration Survey suboperations is subject to CPEX study.

Census Review Analysis and Visualization Application (CRAVA): Data visualization and analytical software developed by the Census Bureau to permit its staff and subject matter experts to review and correct problems during the post data collection processing stage of the census (i.e., progressing from *Decennial Response File* to *Census Unedited File* to *Census Edited File*). It is no longer in use.

Census Questionnaire Assistance (CQA): In the 2020 Census, the operation responsible for telephone-based operations in the census. As the name implies, a main focus of the operation was a dial-in line through which census respondents could obtain information about completing the census form and request a paper questionnaire or a foreign language questionnaire/language assistance guide. Callers to the CQA line were asked if they wanted to complete the census questionnaire during the call; if so, CQA agents conducted the census interview (thus making CQA the source for telephone-based *Self-Response* to the census). In response to the COVID-19 pandemic, CQA functionality was expanded to include some provision for outbound CQA calling (e.g., to try to complete *Nonresponse Followup* cases where physical access was difficult).

Census Unedited File (CUF): The second major step in the product flow for post data collection processing in the 2020 Census, following the *Decennial Response File* 2. The CUF reflects the application of household count imputation and so represents the final universe of address and population counts; accordingly, the state-level *apportionment* counts are generated from the CUF.

CFSA: *Census Field Supervisor Area*

CNSTAT: Committee on National Statistics of the National Academies of Sciences, Engineering, and Medicine

Count Question Resolution (CQR): An operation in the 2020 Census and other recent censuses in which state, local, and tribal governments can challenge their population totals if the officials can demonstrate flaws in the geographic boundaries used to generate population counts or in the location (or associated population count) of housing units. It is a postcensus review activity focused only on housing units, and so differs in scope from the new-in-2020 *Post Census Group Quarters Review* program; it also differs from the *Count Review* program in the late stages of conducting the census, which engages state-level participants to review housing unit counts for general accuracy.

Count Review (CR): Late-stage census operation in which state-level members of the *Federal-State Cooperative for Population Estimation* are permitted to review preliminary counts near the end of data collection, to suggest necessary updates or revisions to the *Master Address File*.

Copyright National Academy of Sciences. All rights reserved.

Coverage Correction Factor (CCF): See *Coverage Measurement*.

Coverage Measurement:[1] A primary means of measuring census quality by estimating undercount and overcount for the nation as a whole and for specific demographic groups. The set of census enumerations from a sample of clusters of census blocks (known as the *E-sample*) is matched to the results of the independent *Post-Enumeration Survey* (PES) conducted in the same sample of block clusters, and the statistical technique of *dual-systems estimation* is used to estimate under- or overcount. In brief:

- Every enumeration recorded in the census (and hence the E-sample) is either "data-defined" (containing sufficient information to enter census processing and editing) or a "whole-person imputation" (in which every person characteristic had to be imputed).
- Enumerations that are not whole-person imputations may be divided into two groups, "correct enumerations" and "erroneous enumerations," with the latter principally consisting of duplicate entries.
- In this way, the major components of census coverage are correct enumerations, erroneous enumerations, whole-person imputations, and omissions.
- Initial Coverage Correction Factors (CCFs) for a particular population group are obtained by multiplying the estimated probability that a census enumeration is data defined by the estimated probability that the enumeration is correct, and dividing by the estimated probability that the enumeration is a match between the E- and P-samples. (Some CCFs are further adjusted based on *Demographic Analysis* sex ratios to account for known, systematic effects.) The sum of the CCFs for the population group is a synthetic estimate of the size, which is then compared to the census count to estimate net coverage rate

CPEX: *Census Program of Evaluations and Experiments*
CQA: *Census Questionnaire Assistance*
CQR: *Count Question Resolution*
CR: *Count Review*
CRAVA: *Census Review Analysis and Visualization Application*
CSAC: Census Scientific Advisory Committee
CUF: *Census Unedited File*
DA: *Demographic Analysis*
DAS: *Disclosure Avoidance System*

[1]In the 2020 Census, Coverage Measurement is the heading used to describe component operations related to this work, though the underlying survey (and hence the program as a whole) is commonly termed the Post-Enumeration Survey (PES). In the 2010 Census, the corresponding suite of operations was known as Census Coverage Measurement (CCM); in the 2000 Census, it was the Accuracy and Coverage Evaluation (A.C.E.).

Copyright National Academy of Sciences. All rights reserved.

DDHC; DDHC-A; DDHC-B: Detailed *Demographic and Housing Characteristics Files*

DDP: *2010 Demonstration Data Products*

Decennial Response File (DRF): The first major step in the product flow for post data collection processing in the 2020 Census, in two stages. The Decennial Response File 1 (DRF1) is the general compilation of all responses by all modes, including duplicate responses; Decennial Response File 2 (DRF2) reflects unduplication of census returns across households using the Census Bureau's proprietary Primary Selection Algorithm.

Delivery Sequence File (DSF): The master database of deliverable mail addresses maintained by the U.S. Postal Service, organized by carrier route. First used as a source of address updates to the Census Bureau's *Master Address File* for the 2000 Census, regular "refreshes" from the DSF remained a primary source of new address information in the 2010 and 2020 Censuses.

Demographic Analysis (DA): A census coverage measurement technique that derives population estimates of national and major demographic subgroups (that are independent of the current census) based on *administrative records* data on births, deaths, immigration, and emigration. Demographic Analysis estimates have been extended back to the 1940 Census.

Demographic and Housing Characteristics Files (DHC): Series of 2020 Census data products largely replacing the content of previous decades' Summary File releases.

DHC: *Demographic and Housing Characteristics Files*

Differential privacy (DP): A system for sharing information that imposes quantifiable limits on the disclosure of information about individual records in the database, as through the injection of statistical noise into all data cell entries. In the 2020 Census, the Census Bureau decided to implement a new *Disclosure Avoidance System (DAS)* premised on differential privacy theory; the *TopDown Algorithm* is the specific differentially private (or formally private) algorithm used to generate the first major 2020 Census data products after the state-level *apportionment* populations: namely, the redistricting data required under P.L. 94-171 and the forthcoming *Demographic and Housing Characteristics Files*. See also *privacy-loss budget*.

Disclosure Avoidance System (DAS): The intermediary system between the *Census Edited File* and the final data files used for tabulation of census products that applies techniques to curb the risk of identification of census respondents. For the 2020 Census redistricting data, the 2020 Census DAS uses the *TopDown Algorithm* to take *differentially private* "noisy measurements"—crosstabulation cells with a controlled amount of statistical noise added to them—and convert them into a synthetic Microdata Detail File that is used for tabulation purposes.

Copyright National Academy of Sciences. All rights reserved.

DOB: date of birth
DP: *differential privacy*
DRF; DRF1; DRF2: *Decennial Response File*
DSE: *dual-systems estimation*
DSF: *Delivery Sequence File*
Dual-systems estimation (DSE): Statistical procedure for estimating the size of a population by combining information from two independent listings/samples of that population. See *coverage measurement*.
E-sample: In *dual-systems estimation*, the set of decennial census (enumeration) returns for the same set of blocks in which the independent *Post-Enumeration Survey* is conducted.
ECaSE-ISR: Enterprise Census and Surveys Enabling–Internet Self-Response
Enumeration at Transitory Locations (ETL): The operation in the 2020 Census and other recent censuses that counts the population at sites where the location itself or residence in individual units is inherently transitory or impermanent (that is, where people are unlikely to live year-round). A key characteristic of transitory locations is that people either pay a fee to stay there (hence the 2020 Census transitory location categories of campgrounds, recreational vehicle (RV) parks, marinas, and hotels/motels) or work there on an impermanent basis (hence, the categories of racetracks and circuses/carnivals). Transitory locations are sometimes confused with *group quarters*, and there is also potential overlap between transitory units (individual units or spaces in transitory locations) and housing units, with an occupied transitory unit being considered a housing unit if at least one person usually lives there. The ETL operation is preceded by a Transitory Location Advance Check operation to broker arrangements for the count.
EPIK: enhanced *protected identification key*
ϵ (epsilon): See *privacy-loss budget*.
eResponse: Response mode in the 2020 Census *Group Quarters Enumeration* operation through which *group quarters* facility administrators could enter response data for all residents as of Census Day in a Microsoft Excel template provided by the Census Bureau and submit it through a secure internet portal.
Erroneous enumeration: See *coverage measurement*.
ETL: *Enumeration at Transitory Locations*
FCSM: Federal Committee on Statistical Methodology
FERPA: Family Education Rights and Privacy Act
Field Verification (FV): Field operation that occurs as part of *Nonresponse Followup* and uses the same enumerators, but in which the focus is solely on verification of housing unit addresses newly detected during census data collection and collection of geographic coordinates for the address (not the conduct of an interview with household respondents). In the 2020 Census, FV cases were those addresses originating from *Non-ID Processing* that

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

could not be automatically matched to the *Master Address File* or resolved in *In-Office Address Verification*.

FSCPE: Federal-State Cooperative for Population Estimates

FV: *Field Verification*

FY: fiscal year

GAO: U.S. Government Accountability Office

GQ: *group quarters*

GQAC: *Group Quarters Advance Visit*

GQE: *Group Quarters Enumeration*

Group quarters (GQ): A place where people live that is not a housing unit. Group quarters are commonly categorized as institutional (e.g., correctional facilities and nursing or other health care facilities) or non-institutional (e.g., college student housing, military barracks, ships, group homes, worker dormitories, and shelters). GQs differ from housing units because GQs are owned or managed by some entity providing housing or custodial/medical services to the residents; they differ from *transitory locations* because people commonly pay a fee for an impermanent stay at a transitory location.

Group Quarters Advance Visit (GQAC): The census field operation through which initial contact is made with a *group quarters* site, basic information is collected to properly categorize the site's group quarters type, and arrangements are made for data collection in the *Group Quarters Enumeration* operation.

Group Quarters Enumeration (GQE): The census field operation through which data is collected from or about *group quarters* residents, whether through administration of individual questionnaires or provision of data from the facility as a whole (e.g., in the 2020 Census, through the *eResponse* internet response option for group quarters). Emergency and transitional shelters, soup kitchen and regularly scheduled mobile food van locations, and *targeted non-sheltered outdoor locations* are all considered group quarters in the decennial census framework, and so the *Service-Based Enumeration* operation that covers those locations that may be used by people experiencing homelessness is an operational adjunct to the broader GQE.

GSS-I: Geographic Support Systems Initiative

HUD: U.S. Department of Housing and Urban Development

IFAC: *In-Field Address Canvassing*

In-Field Address Canvassing (IFAC): See *Address Canvassing*.

In-Office Address Canvassing (IOAC): See *Address Canvassing*; *Active Block Resolution*; *Interactive Review*.

INR: item nonresponse

iCADE: The Integrated Computer Assisted Data Entry system to scan paper questionnaires and capture the data in electronic format for further

Copyright National Academy of Sciences. All rights reserved.

processing. The iCADE system was developed within the Census Bureau and has been used in various Census Bureau surveys and some international censuses. Paper data capture was outsourced in the paper-based 2010 Census, but the in-house iCADE system was adopted to process paper questionnaires in the 2020 Census.

Interactive Review: Lead portion of the 2020 Census In-Office *Address Canvassing* program, beginning in fall 2015 and continuing through early 2019, in which analysts viewed *Master Address File* housing unit locations and counts superimposed on aerial imagery to identify change or stability in residential structures. Interactive Review was meant as a comprehensive but quick triage, classifying blocks as active (evidence that the number of housing units visible in imagery differs from the number in the Master Address File), passive (stable blocks, in terms of housing unit numbers in both sources), and on hold (blocks for which the available imagery was inconclusive). In the original program, active blocks were sent to the *Active Block Resolution* stage; after that stage was discontinued, blocks unresolved in Interactive Review were sent to In-Field Address Canvassing (though a number of "trigger" events, such as the availability of updated imagery, could prompt another round of Interactive Review for blocks).

Internet Choice; Internet First: The two principal mail contact strategies assigned to census tracts in the *Self-Response type of enumeration area* based on estimates of likelihood to respond via the internet, as derived from American Community Survey (ACS) data. Internet Choice areas, with older populations and low internet take-up in the ACS, received a paper questionnaire in the initial mailing (along with instructions for responding online); Internet First areas, deemed more likely to readily respond via the internet, received multiple letter/postcard cues to respond online before a paper questionnaire was sent to still-nonresponding households.

Internet Self-Response (ISR): A designated operation of the 2020 Census to facilitate online response to the census. Core to the operation was the development of an internet response platform—first the Census Bureau-developed *Primus* system, then the outsourced *eCaSE-ISR* system, and ultimately a return to Primus.

IOAC: *In-Office Address Canvassing*

IRS: Internal Revenue Service

IT: information technology

JASON: An independent group of scientists engaged to advise the federal government on science and technology matters, principally in the defense and intelligence arenas. JASON administratively operates through the MITRE Corporation; despite the capitalization, JASON is not an acronym.

Local Update of Census Addresses (LUCA): A Census Bureau program developed in response to the Address List Improvement Act of 1994 (and so included in the 2000–2020 Censuses) in which local officials were given

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

the opportunity to review individual addresses on the *Master Address File* and make corrections, additions, and deletions to that list, and to make corrections to census maps to match any changes that may be needed.

LUCA: *Local Update of Census Addresses*

MAF: *Master Address File*

Master Address File (MAF): The master list of residential and nonresidential addresses maintained by the Census Bureau. Since the 2000 Census, the MAF has been maintained as an ongoing resource for the Census Bureau's censuses and surveys, updated using a variety of sources including the U.S. Postal Service's *Delivery Sequence File* and, in the decade preceding the 2020 Census, a major Geographic Support System initiative to obtain address and geographic files from state, local, and tribal governments. MAF entries are linked to a specific geographic location (for eventual tabulation purposes) using the Census Bureau's principal cartographic resource, the Topologically Integrated Geographic Encoding and Referencing (TIGER) System. Extracts are derived from the MAF using a series of filters to support specific operational purposes; the report text examines two of these extracts, denoted MAF1 and MAF2, that correspond to the set of addresses deemed workable for enumeration at the beginning of 2020 Census operations and the set of addresses enumerated in the 2020 Census, respectively.

MAFID: Master Address File ID (identifier)

MAPE: mean absolute percentage error

MENA: *Middle Eastern or North African*

Metadata: Data about other data, including quality metrics data (e.g., item nonresponse rates), data on the progress of individual operations in a census or survey, or *paradata* concerning the manner of data collection.

Middle Eastern or North African (MENA): An additional major race category that has periodically been suggested for addition to the federal standards in *Statistical Policy Directive No. 15*.

Mobile Questionnaire Assistance (MQA): An adjunct to the 2020 Census *Internet Self-Response* operation, added very late in the planning process in response to Congressional appropriations language. MQA was intended to deploy temporary Questionnaire Assistance Centers at selected, large public gatherings in areas with low self-response rates. Program staff with Census Bureau-issued mobile devices would allow visitors to complete their census form online. However, the operation was largely scaled back due to COVID-19-related prohibitions on public gatherings.

Modified Race File: A special tabulation of census data that allocates *Some Other Race* Alone responses (i.e., in which the only major race category selected was Some Other Race) to other specific race groups, based on the ethnicity question and other factors. This file is necessary for detailed *Demographic Analysis* results by race because vital records and other data

Copyright National Academy of Sciences. All rights reserved.

sources do not include Some Other Race as an option; production of the Modified Race File has been greatly slowed in the 2020 Census for lack of a disclosure avoidance routine consistent with the *differential privacy*-based 2020 *Disclosure Avoidance System*.

MQA: *Mobile Questionnaire Assistance*

MWI: Modified Whipple's Index

National Processing Center (NPC): The Census Bureau's principal facility for collecting and processing the returns from its censuses and surveys, as well as for providing logistical and clerical support for decennial census field operations. The NPC is located in Jeffersonville, Indiana. In the 2010 Census, the NPC housed a data capture center that shared paper data capture responsibilities with two other centers (in Essex, Maryland, and Phoenix, Arizona) that were administered by external contractors; in the 2020 Census, paper data capture was shared between the NPC and a facility in Phoenix.

NC: *New Construction*

New Construction (NC): A census geographic partnership and address-building operation in which state, local, and tribal governments submit new mailing addresses for living quarters under construction but that were expected to be complete by Census Day.

Non-ID Processing: The system in the 2020 Census to facilitate the collection of respondent information (including self-supplied usual residence information) without requiring the respondent to enter a *Master Address File*-keyed identifier code from a physical mailing to a household. In the 2010 Census, Non-ID nonreturns were primarily the blank-address "Be Counted" forms made available in public places for people to complete if they believed they had been missed in the census. However, in the 2020 Census, Non-ID Processing was critical to the extensive promotion of the notion of replying to the census anytime or anywhere, using the Internet Self-Response option. The address information collected in Non-ID Processing is compared and reconciled with the Master Address File, with ungeocoded addresses being designated for verification by field staff.

Nonresponse Followup (NRFU): The census field operation whereby census enumerators attempt to collect responses from households for which no response was obtained by Self-Response methods (in the 2020 Census, by Internet, mail, or phone) or returning a questionnaire left with the housing unit in *Update Leave*. In the 2010 Census, NRFU was performed exclusively using paper questionnaires, with assignments parceled out to individual enumerators through a hierarchy of field staff. The 2020 Census utilized electronic interviewing on smartphones, with assignments pushed out to enumerators' devices on a nightly basis; particularly in the early stages, the enumerator was also given an optimized route to follow in conducting the interview attempts. NRFU in the 2020 Census was also

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

the first in which the availability of reliable *administrative records* data for a household could be used for enumeration in lieu of second and subsequent enumerator visits.

NPC: *National Processing Center*

NRFU: *Nonresponse Followup*

NRFU RI: *Nonresponse Followup Reinterview*; a quality assurance operation

Numident: Short for Numerical Identification System, the database maintained by the Social Security Administration for assignment of Social Security numbers and registration in the system.

OBAV: *Office-Based Address Verification*

Office-Based Address Verification (OBAV): An operation in the 2020 Census in which Census Bureau office staff worked to resolve addresses coming in through *Non-ID* returns (i.e., without reference to the printed identifier number in any census mailing or delivery) that did not readily match to the *Master Address File*, without requiring a field visit.

OMB: U.S. Office of Management and Budget

P-sample: Set of returns from the *Post-Enumeration Survey*, a sample survey of respondents in a sample of block clusters conducted after (and independent of) the census enumeration; part of the Census Coverage Measurement program, it is used to calculate the match rate in the *dual-systems estimation* formula.

Paradata: Operational or functional data from a census or survey that describe the manner in which the census/survey data were collected, such as the number of contacts made with a survey respondent or the mode of data collection. Paradata is essentially a subset of *metadata*.

PCGQR: *Post Census Group Quarters Review*

Person Identification Validation System (PVS): The Census Bureau computer system for assigning *protected identification keys* to *administrative records* and other data sources to facilitate data linkage.

PES: *Post-Enumeration Survey*

PIK: *protected identification key*

Post Census Group Quarters Review (PCGQR): A special review program permitting state and local governments to challenge and correct the 2020 Census *group quarters* enumeration in their jurisdictions; though the results do not change the official 2020 Census totals (for redistricting and other purposes), it allows for correction of annual population estimates going forward.

Post-Enumeration Survey (PES): Both a generic term for the independent, follow-up survey conducted in census *coverage measurement* programs based on dual-systems estimation and the specific operational name given to said survey in some decennial censuses (as done in 2020).

Primus: The technical platform for Internet Self-Response in the 2020 Census. Primus was developed in-house by the Census Bureau in early 2020 Census

Copyright National Academy of Sciences. All rights reserved.

planning and testing but was shifted to backup-system status in mid-decade in favor of an externally developed software solution. Immediately prior to the 2020 Census, the Census Bureau concluded that Primus was better able to handle high volumes of concurrent internet users, returning it to use as the primary response channel in the census.

Privacy-loss budget: The parameter in a *differential privacy* implementation, technically defined as the maximum distance between queries performed on two databases that differ from each other by one entry, that is conceptualized as the amount of information disclosure risk that will be tolerated. The privacy-loss budget is commonly denoted by ϵ and characterized as a slider between privacy and accuracy, $\epsilon \rightarrow 0$ meaning perfect privacy and zero accuracy and $\epsilon \rightarrow \infty$ connoting absolute accuracy and zero privacy. The *Disclosure Avoidance System* allocates certain shares of overall ϵ to specific histogram cells and levels of geographic tabulation, thus determining the amount of statistical noise added to the cells.

Protected identification key (PIK): The anonymized identifier variable assigned to entries in *administrative records* source files and other data resources to enable record matching across sources.

Proxy: A census interview in which the respondent is not a member of the household being enumerated. The respondent might be a neighbor, a landlord, or some other knowledgeable person.

PVS: *Person Identification Validation System*

Regional Census Center (RCC): See *Area Census Office*.

ROCkIT: Re-Organize Census with Integrated Technology

SafeTab-H; SafeTab-P; PHSafe: *Differential privacy*-based algorithms developed for generating the Detailed *Demographic and Housing Characteristics* Files and other late-stage 2020 Census data products.

SAIPE: Small-Area Income and Poverty Estimates

SBE: *Service-Based Enumeration*

Service-Based Enumeration (SBE): In the 2010 and 2020 Censuses, the component of the broader *Group Quarters* Enumeration program focused on people experiencing homelessness who could be reached at service or shelter locations or in certain *targeted non-sheltered outdoor locations*. In 2010, the program was spread across three days around the April 1 reference date, each day specific to a type of location; in the 2020 Census, SBE was postponed until September 22–24.

SF1; SF2: *Summary File 1; Summary File 2*

Some Other Race (SOR): The sixth major race category permitted by the Census Bureau for response to the race question. It is not defined or required by the U.S. Office of Management and Budget's *Statistical Policy Directive No. 15* but is rather required by Congress through language in Census Bureau appropriations.

SOR: *Some Other Race*

Copyright National Academy of Sciences. All rights reserved.

SRQA: Self-Response Quality Assurance
SSN: Social Security number
Statistical Policy Directive No. 15: Common name for the 1977 version of standards issued by the U.S. Office of Management and Budget for the collection of data on race and ethnicity (Hispanic origin). The most recent revision (1997) of these guidelines is not labeled or numbered as a directive or circular. The guidelines make it possible to identify with more than one race, rather than choosing from only one of a few mutually exclusive race categories.
Summary File 1 (SF1); Summary File 2 (SF2): The detailed data files of tabulations that were the primary data products of recent U.S. censuses, until replaced in the 2020 Census by the *Demographic and Housing Characteristics Files*.
Targeted non-sheltered outdoor locations (TNSOL): Technically designated as noninstitutional *group quarters* locations, areas where people experiencing homelessness are known to live without paying to stay, and where they are counted in the *Service-Based Enumeration* program.
TDA: *TopDown Algorithm*
TEA: *type of enumeration area*
Tenure: The status of an occupied housing unit as either owner occupied or renter occupied.
TIGER: Topologically Integrated Geographic Encoding and Referencing (TIGER) System; the Census Bureau's framework for identifying the exact geographic location of residential addresses (as well as other physical features).
TL: transitory location; see *Enumeration at Transitory Locations*
TNSOL: *targeted non-sheltered outdoor location*
TopDown Algorithm (TDA): The *differential privacy*-based algorithm developed by the Census Bureau to produce the P.L. 94-171 Redistricting Data File and *Demographic and Housing Characteristics Files*. Generally speaking, the algorithm works separately with person and housing unit data from the *Census Edited File*, injecting a specified amount of statistical noise into every entry of summary tables/queries (based on allocation of *privacy-loss budget* to each query) and using nonlinear optimization and other tools on the resulting "noisy measurements" to find a microdata representation that satisfies both noise-injected table values and additional constraints (such as nonnegative counts or totals that must remain invariant to change). The name arises from the manner in which the algorithm generates national-level estimates first, then iterates to produce estimates at the state level, county level, and so forth, down to the census block level.
Transitory location (TL): See *Enumeration at Transitory Locations*
Type of enumeration area (TEA): A classification developed by the Census Bureau prior to a decennial census to determine the basic strategy used to

Copyright National Academy of Sciences. All rights reserved.

conduct the census in particular blocks (or, in 2020 Census terminology, basic collection units). The 2000 and 2010 Censuses each designated nine TEAs, while the 2020 Census utilized six TEAs: Self-Response (subsuming the Mailout/Mailback TEA of the previous censuses), *Update Enumerate*, *Update Leave*, Island Areas (separate designation for the U.S. territories), Remote Alaska, and Military.

UAA: *Undeliverable as Addressed*

UE; UE/RA: *Update Enumerate; Update Enumerate/Remote Alaska*

UL: *Update Leave*

Undeliverable as Addressed (UAA): U.S. Postal Service terminology for any mailing that cannot be delivered to the specific name and address on the mailing, for any reason.

Update Enumerate: A method of census enumeration and *type of enumeration area* in which a census enumerator's sole visit involves both *Address Canvassing* functions (adding or updating address list entries) and enumeration (conducting an interview with household respondents). Update Enumerate areas tend to be remote, rural geographic areas; Remote Alaska areas are counted using Update Enumerate methods but are designated as a separate type of enumeration area because they are approached earlier in the census year (beginning in January) due to physical accessibility concerns.

Update Leave: A method of census enumeration and *type of enumeration area*, primarily focused on geographic areas where the U.S. Postal Service does not do home delivery or in which there are other possible issues with mail delivery, in which a census enumerator conducts address list update activities in the area and leaves a questionnaire package at each housing unit. Update Leave was not originally planned to be a type of enumeration area in the 2020 Census, as it was in 2010, but it was reinstated in the 2020 Census plan in 2017. In addition to areas lacking mail delivery, Update Leave has also been applied to areas that have been affected by major disasters prior to the census.

USPS: U.S. Postal Service

Whole-person imputation: A census enumeration for which all characteristics information is unknown and requires imputation, occurring when a household is believed to be occupied but no response data of any kind is available.

zCDP: Zero-concentrated *differential privacy*; a variant of pure differential privacy that relaxes some assumptions. Beginning April 2021, the Census Bureau's *TopDown Algorithm* was reconfigured to use discrete Gaussian noise infusion and zCDP in lieu of discrete Laplace noise infusion and pure differential privacy.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

– C –

Additional Detail and Reference on 2020 Census Operations

This appendix summarizes procedures used to conduct the 2020 Census, based primarily on U.S. Census Bureau presentations made in public sessions of the panel in 2021 (see Appendix B in the interim report). It has sections on the following: (1) goals for the 2020 Census; (2) hiring, training, and retention of field staff; (3) development of the Master Address File (MAF) used to initially contact households through the U.S. Postal Service (USPS) or delivery by census field staff; (4) types of enumeration areas (TEAs); (5) attempts to encourage self-response by household residents through outreach, partnerships, and promotion; (6) contacting household residents to provide multiple ways of responding and enumerating people in group quarters (GQs) and other nonhousehold situations; (7) attempts at following up initial nonrespondents and closing out nonresponsive households; (8) data processing and creation of data products; and (9) assessments of the quality of the counts and the effectiveness of various census processes.

C.1 GOALS FOR THE 2020 CENSUS

The processes used to conduct the 2020 Census were designed to achieve the following goals: (1) to provide a complete and accurate count so that every resident is counted once, only once, and in the right location; (2) to contain costs and improve efficiency in field operations; (3) to maintain security of data during transmission and on the Census Bureau's computing network; and (4) to

Copyright National Academy of Sciences. All rights reserved.

provide alternative ways for the public to self-respond. A goal added toward the end of 2020 Census planning was to use newly developed disclosure avoidance techniques for the 2020 data products.

To achieve these goals in a cost-effective manner, an initial objective of the 2020 Census was to reengineer field operations. Field reengineering had three objectives: (1) streamline the office and staffing structure; (2) increase the use of technology; and (3) increase the effectiveness of field management. Steps to improve field management included giving managers tools to provide better data on the status of cases in the field, along with improved communications and redesigned quality-assurance operations. Technology was used to automate the provision of work assignments for field personnel, automate aspects of field personnel recruitment and training, automate compensation and expense reporting, and reduce reliance on paper and the manual processing of responses.

C.2 HIRING, TRAINING, AND RETENTION OF FIELD STAFF

Field staff were essential to many 2020 Census operations, including: In-Field Address Canvassing, Remote Alaska, Update Leave, Update Enumerate, some Group Quarters Enumeration operations, and, most importantly, Nonresponse Followup (NRFU). The Census Bureau implemented a web-based Recruitment and Assessment System—over 3.9 million applicant profiles were created in the system over the course of the 2020 Census. Almost 3.1 million applicants were available for job selection. The Census Bureau employed over 500,000 temporary staff during the taking of the census, primarily for NRFU, with nearly 300,000 people working during the peak week of NRFU in early August 2020.

Because of COVID-19 restrictions, field training for NRFU had to be modified from a partially office-based approach to a virtual training program. There were also 2 hours of in-person enumerator training during which enumerators received equipment. (The original plan had been for 2 days of in-person training.) In addition, extraordinary efforts were made to retain field staff. The Census Bureau extended the recruitment period by 5 months to run through September 2020 and offered incentives to enumerators for greater production. High-performing enumerators who agreed to travel were asked to help complete NRFU in areas that were experiencing production shortfalls.

Field operations were facilitated by technology. Address canvassers used laptops equipped with the Listing and Mapping Application (LiMA) to complete in-field address work. NRFU enumerators used smartphones. Assignments were generated online and adjusted dynamically—for example, to account for late self-responses in the NRFU assignments.

Copyright National Academy of Sciences. All rights reserved.

C.3 CREATING THE ADDRESS LIST FOR THE 2020 CENSUS

As described in more detail in Chapter 5, the 2020 Census continued the practice of using an ongoing repository of address information—the MAF—as the basis for the operational frame for the census. It differed slightly in terminology; in the 2000 and 2010 Censuses, the notion was to construct a snapshot extract immediately prior to the start of main census operations, which was termed the Decennial MAF (DMAF), while the 2020 Census was couched as applying various filters to the ongoing, fluid MAF database to define specific operational universes. But the MAF-building process for 2020 differed greatly in methodology and approach, through the major innovation of in-office canvassing methods to review the MAF addresses without fieldwork (and, pointedly, without a 100-percent field canvass of every block in the nation as was done in 2000 and 2010).

C.3.1 2010 Decennial Master Address File Development

An overview of the analogous process used for the development of the 2010 DMAF is offered for purposes of comparison. The goal of the 2010 DMAF development process was to provide the primary address list update and validation activity for the 2010 Census enumeration frame. The Census Bureau deployed a national-level, automated, paperless data-collection and transmission process, primarily making use of handheld computers that were provided to census field workers. The Address Canvassing operation, which covered 100% of the residential addresses in the United States, was managed by opening 150 Local Census Offices in 12 Regional Census Centers early. It was conducted by over 150,000 field staff who started on March 30, 2009, and who completed their work on July 10, 2009.

In addition to verifying addresses for census mailing purposes, a major focus of the 2010 Address Canvassing was the collection of GPS coordinates (a "map spot") for every entry on the MAF, for inclusion in the MAF and the Census Bureau's companion Topologically Integrated Geographic Encoding and Referencing (TIGER) System and to associate each MAF entry with a specific geographic location.[1] The exception was large blocks that contained more than 1,000 addresses, which were handled by the Large Block Address Canvassing (LBAC) operation, which was conducted from February 2, 2009, until June 17, 2009. The LBAC used laptop computers to collect information on the addresses because the handhelds had limited storage space.

[1]Prior to the collection of map spot data—or in the absence of those coordinates—geographic locations and placement of address numbers within census blocks can be done by exploiting the topological nature of the TIGER database. Line segments corresponding to streets are also coded with beginning and end addresses, on both sides of the street as applicable, so a specific street address can be assigned an approximate geographic location by interpolation along the line segment.

Copyright National Academy of Sciences. All rights reserved.

For GQs—primarily college/university student housing, residential treatment centers, skilled nursing facilities, group homes, correctional facilities, and workers' dormitories—the 2010 address canvassing operation identified each possible residence as either a housing unit, a GQ that was included in the 2000 Census DMAF inventory, or as a potential GQ from other sources, such as administrative records. Address canvass listers marked addresses that were not housing units as Other Living Quarters (OLQs); those addresses were included in the 2010 Census Group Quarters Validation (GQV) operation universe. Specially trained staff used the GQV questionnaire to verify that the address had the correct census geography and to determine the status of the OLQ address as either a GQ, a housing unit, a transitory location, or as nonresidential, vacant, or nonexistent. Staff also classified the type of GQ and collected the maximum number of residents who could live or stay at the address. These additional GQs, housing units, and transitory locations were added to the developing decennial master address list.

C.3.2 Intercensal Additions to the MAF

Master Address File Coverage Study

In previous decennial censuses, the Census Bureau used a 100% block canvass to update the MAF from the previous census. This operation was increasingly viewed as inefficient, since a large percentage of residential blocks experience no change from one census to the next. As a result of this finding, the development of the MAF for the 2020 Census was planned to be conducted in two phases. There would be an "in-office" phase, conducted by census geographers who used imagery to identify stable blocks versus "active" blocks that experienced changes from the previous census (i.e., differed from the current MAF). After the "in-office" phase, there was an "in-field" phase, conducted by field personnel, carried out as a targeted block canvass on the subset of residential blocks that were identified as being subject to change in the intercensal period. This canvassing attempted to physically locate each address and determine whether it was a housing unit.

The MAF Coverage Study was intended to be an ongoing field activity that listed a sample of 20,000 nationally representative blocks annually. This sample was planned to measure the coverage of the MAF, validate the use of in-office canvassing procedures that could determine which blocks had additions or deletions during the intercensal period, improve in-field data-collection methods, and provide updates to the MAF on a continuous basis. Field work started in April 2016, but only one report was produced—the June 2016 Monthly Status Report—and the program was discontinued in 2017 due to budgetary constraints.

Copyright National Academy of Sciences. All rights reserved.

U.S. Postal Service Delivery Sequence File Updates to the Master Address File

Throughout the decade of 2010–2020, various versions of the USPS's Delivery Sequence File (DSF) were used to update the ongoing MAF. The USPS provided three files to the Census Bureau twice a year during the intercensal period: (1) the DSF, a database of all mail delivery addresses that, as implied by the name, is used to structure routes for mail carriers; (2) a locatable address conversion service file; and (3) a list of ZIP codes and the plus-four codes served by USPS. MAF/TIGER uses these addresses and locations to update the information on current addresses and to add new addresses to the MAF in what is often termed a "refresh." The version of the MAF for use in the subsequent census, the DMAF, is initiated through updating about a year prior to the start of the census using the most current DSF. That updating, in 2018, added 5.3 million new residential addresses to the final version of the DMAF from the 2010 Census. Another 2.3 million current residential addresses were in the 2010 DMAF alone. While there is substantial overlap, both systems—DMAF and DSF—have unique addresses.

Updates to the Master Address File through Partnerships

Geographic Support System (GSS) Beginning in 2011, the Census Bureau implemented a Geographic Support System Initiative for continuous updating of MAF/TIGER and related activities, working with public- and private-sector partners. The initiative became the Geographic Support System (GSS) Program in 2015. Its activities included: maintaining the MAF/TIGER system; collecting and assigning addresses to geographic locations; identifying, collecting, and maintaining geographic entities and their boundaries; preparing and maintaining maps that support data collection and dissemination of Census Bureau data; developing and maintaining partnerships with tribal, federal, state, and local governments and commercial entities for the acquisition of address and mapping data; and coordinating federal and international geospatial activities with other federal agencies and the appropriate international groups. The GSS was coordinated with the Local Update of Census Addresses (LUCA) Program in 2018, attempting to acquire address or road data for 15,000 census tracts located within government entities that were not participating in LUCA.

Through the GSS, 106.7 million addresses were acquired between 2012–2018. Of those addresses, 106.2 million (99.5%) matched addresses already in the 2010 DMAF, which was confirmation of its quality. Geospatial locations were improved or corrected for 75.1 million addresses. In addition, through the Ungeocoded Resolution Project, 810,899 addresses (72% of addresses reviewed) that previously were not assigned to a census block were geocoded.

Copyright National Academy of Sciences. All rights reserved.

Local Update of Census Addresses (LUCA) LUCA resulted from the Census Address List Improvement Act of 1994. The Act allowed tribal, state, and local governments to review and provide updates to the Census Bureau's MAF for their jurisdictions under the condition of confidentiality.[2] Updates could include addresses that were missed, addresses that were included in error, and addresses located in the wrong block. This updating began two years prior to the decennial census to allow for completion of an appeals process to finish (LUCA Validation). The LUCA operation for the 2020 Census began in April 2018 and was completed in March 2019.

LUCA consisted of the following phases: (1) *LUCA Outreach*—promotions, training, and registration; (2) *LUCA Review of Materials*—submission of address lists by participants; (3) *LUCA Returns and Validation*—changes made to the MAF by the Census Bureau; and (4) *LUCA Feedback and Appeals* and *LUCA Closeout*. LUCA validation included a full in-office review of LUCA-submitted changes that were not validated through an automated match to the MAF. Such changes included potential additions that did not match an existing GSS or MAF record, addresses that were moved by the local participant to a different block than indicated in MAF/TIGER, and additions of nonresidential addresses to the MAF. Of the addresses provided, 81.2% were matched to the current version of the MAF, leaving 3.46 million unmatched addresses.

In-Office Address Canvassing

In-Office Address Canvassing (IOAC), with initial review of the country completed in June of 2017, and the entire operation completed by October 11, 2019, was based on satellite imagery used to detect changes or stability in blocks of housing units throughout residential areas in the United States and Puerto Rico. This program compared the number of housing units identified in the imagery to the number in the Census Bureau's MAF/TIGER system. IOAC included Interactive Review, Active Block Resolution, Ungeocoded Resolution, and Group Quarters/Transitory Locations review. It also reviewed LUCA addresses that failed automatic matching to the MAF.

Interactive Review (IR) IOAC used imagery and validated data sources to identify changes and potential changes (i.e., new construction) on the ground, related to residential structures. The intent was to eliminate field work in areas with no measured changes or where the residential structure inventory was stable. Office staff used validated data sources to correct any identified issues and to resolve addresses that could not be automatically geocoded. Staff measured the extent to which the number of housing units in the MAF was

[2]In contrast, GSS was a one-way sharing of addresses from partners to the Census Bureau and not the reverse.

Copyright National Academy of Sciences. All rights reserved.

consistent with the number of residential structures visible in current imagery. The result was the assignment of blocks to one of three categories: (1) *passive*—no observable change in the number of housing units over time, and the MAF accurately represented the number of housing units shown in the imagery; (2) *active*—observed changes in the residential landscape or the number of housing units in the MAF differed from the number shown in the imagery; or (3) *on hold*—when a review could not be completed due to poor imagery.

Active Block Resolution (ABR) ABR was designed to research active blocks, which showed evidence of change in IR, in the office using street-view imagery and other address resources to attempt to resolve them in the office rather than send the block directly to In-Field Address Canvassing (IFAC). Blocks that were not fully resolved in ABR remained active and progressed to IFAC. ABR launched in April 2016 but was discontinued in February 2017 due to budgetary constraints, leaving IR as the main component of IOAC.

Ungeocoded Resolution (UR) This operation began in April 2017 and finished in February 2020. UR was designed to assign a block location—a geocode—to residential addresses lacking a geocode in the MAF. Staff used data sources in the Matching and Coding System and Geographic Information System viewer. They added or edited spatial features and address ranges in the MAF/TIGER system based on information from local source data.

Group Quarters/Transitory Locations (GQ/TLs) This operation began in September 2017 but stopped in March 2018 due to budgetary constraints. The GQ/TLs operation attempted to verify, update, and validate GQ/TL addresses in the MAF, using administrative data sources, local geographic information systems (GIS) data, public and commercial information, and by calling administrative contacts.

In-Field Address Canvassing

IFAC was similar to the 100% block canvass used in the 2010 decennial census, though it was focused on the subset of blocks for which there was direct evidence of recent change since 2010. In those residential blocks that showed evidence of change in the number of housing units through the decade, listers equipped with laptops were deployed to canvass over 50 million addresses. The resulting effort proved to be 31% more efficient than the initial plan, which called for a 100% block canvass.

Blocks that were determined to need canvassing from IOAC and other sources were:

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

- Blocks identified as having undercoverage during IR, except for those containing only single-family housing and with a history of complete and accurate updates from the USPS DSF within a specified timeframe. Blocks could include undercoverage only or undercoverage in combination with overcoverage, growth, and decline.
- Any block with overcoverage of two or more housing units, and no undercoverage. Blocks could include overcoverage only or overcoverage in combination with growth/decline
- Any block exhibiting growth or decline and an inconsistent history of DSF updates, even when IR did not classify the block as having overcoverage or undercoverage.
- Any block for which IR detected decline but was not classified as having overcoverage or undercoverage.
- Any block with on-hold status, suggesting that the block was not covered by adequate imagery for update, or that future growth could occur.
- Any block that was adjusted through use of the LUCA program.
- Any block that was "triggered" for rereview (e.g., because of a MAF update subsequent to IR) and could not be resolved in-office.

IFAC began on August 4, 2019, and was completed on October 11, 2019. It was conducted in 22% of the Basic Collection Units (BCUs) used in 2020 Census operations,[3] which contained 34.9% of addresses in Self-Response (TEA 1) areas (described below). Address information was collected using laptops managed by the Decennial Device as a Service program. The laptops were equipped with the LiMA to complete work in the field. Field staff canvassed specific geographic areas to identify every place where people could live or stay, comparing what was seen on the ground to the existing MAF, verifying or correcting address and location information, and adding new address records for those missing and removing addresses for residences that did not exist.

The IFAC operation made use of a soft launch, so that the procedures could be beta tested prior to more universal adoption. After IFAC, field staff had to process LUCA appeals and addresses from the New Construction and Count Review operations.

During the 2020 Census itself, several operations provided additional addresses to the MAF. Areas designated for Update Leave (in-person delivery of questionnaire packages) and Update Enumerate/Remote Alaska (in-person interviewing) were ineligible for IFAC because of the "Update" component

[3]The BCU superseded the collection block and assignment area constructs used in the 2010 Census. Collection blocks corresponded in most but not all cases to tabulation blocks, the difference being that collection blocks had to be bounded by visible features and not invisible political boundaries. Assignment areas were groups of collection blocks assigned to field staff to conduct their work. The 2020 BCU does not have a readily available clear definition but combines aspects of the 2020 collection block and assignment area definitions to develop a cluster of addresses that a field worker could cover in a single day. BCUs were adjusted over the course of the decade.

Copyright National Academy of Sciences. All rights reserved.

of those operations; field staff reviewed and updated MAF entries as part of their work in these areas, which were mainly rural and lacking mail delivery. Though address list work was not a primary focus of NRFU, NRFU field interviewers had the capacity to collect MAF information on new addresses that they experienced on their rounds. New for the 2020 Census, the Office-Based Address Validation operation applied IOAC methods to attempt to resolve addresses from Non-ID Processing that did not directly match the MAF. Finally, the special enumeration operations for GQ and transitory locations included an Advance Visit phase that was important for building and updating the address information for those sites.

C.4 TYPES OF ENUMERATION AREAS

Prior to the census, the nation is delineated into TEAs. These assignments, made at fine geographic resolution (collection blocks in 2000 and 2010, BCUs in 2020), govern the way in which initial contacts to participate in the census are made, and also have implications for the way MAF updating proceeds in the areas.

The 2020 Census used five TEAs. The largest of these was TEA 1, Self-Response, which was designed to cover about 95% of the households in the nation. TEA 1 succeeded the Mailout-Mailback TEA designated in the 2000 and 2010 Censuses (the name change reflecting the innovation of internet response), and covered areas with predominantly home mail delivery rather than delivery to post office boxes. TEA 1 was further subdivided, based on American Community Survey (ACS) and other data, into two contact strategy groups (Internet Choice and Internet First, the former receiving a paper questionnaire earlier in the sequence of mailings and contacts than the latter) and into two language groups (with households in some areas with particularly strong concentrations of households with Spanish as the primary spoken language receiving a bilingual English-Spanish questionnaire). The contact strategies are discussed later in this appendix.

The second largest TEA in the 2020 Census was TEA 6, Update Leave,[4] at over 6.8 million addresses, and was applied in areas where the majority of the households do not receive mail at their home's physical location, such as small towns where mail is only delivered to post office boxes or areas recently affected by natural disasters. In these areas, invitation packets that included paper census questionnaires were dropped off at the residences, with the respondents asked to return the filled-in questionnaires either online or via the mail; there was no attempt to complete the interview in person. The "Update" part of the operation's name refers to MAF updating; because census field staff were

[4]In previous censuses, TEAs bore labels punctuated with slashes—Update/Leave and Update/Enumerate—but these were dropped in the early documentation of 2020 Census plans.

Copyright National Academy of Sciences. All rights reserved.

to visit addresses in these areas to deliver questionnaires, TEA 6 areas were excluded from IFAC and the address list update function was made part of the Update Leave work. The reason for Update Leave bearing the numeric label 6 is that it was not part of the original 2020 Census plan—the idea being to handle all primarily rural, non-home-mail-delivery areas as Update Enumerate as described next—but the TEA was reinstated in later versions of the 2020 Census Operational Plan.

The two remaining principal TEAs in the 2020 Census—TEA 2 Update Enumerate and TEA 4 Remote Alaska—collectively comprise less than 1% of households in the nation in mainly remote regions. They share the same methodology: in-person interviewing at the time of the field enumerator's visit, along with the MAF updating implied by the "Update" part of the name. TEA 2 Update Enumerate was used in areas where it was difficult to deliver mail and in areas where internet response was judged to be particularly difficult; in 2020, this includes parts of northern Maine, southeast Alaska, and American Indian and Alaska Native areas that were offered their choice of enumeration procedures. TEA 4, Remote Alaska, as the name implies, covered large portions of Alaska where houses may not be connected by roads and internet connectivity is often poor but was conducted in the same manner of Update Enumerate—with the significant difference of timing. Remote Alaska has become storied and publicized in recent censuses as the earliest field operation, taking place beginning in January 2020 before the spring thaw and before residents leave their villages to fish and hunt. TEA 4 accounted for about 34,000 completed enumerations.

TEA 3, Island Areas, covered American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, and the U.S. Virgin Islands. As with previous decennials, TEA 3 in the 2020 Census defined an essentially parallel census effort: the census was entirely paper-based in these areas and conducted by enumerators who began by listing the addresses and then conducting in-person interviews.

Early plans—and the initial delineation of TEAs for the 2020 Census—included TEA 5, Military, covering areas in and around military installations in the United States. Early plans were to rely on administrative records-type data provided by the U.S. Department of Defense to count the service personnel in TEA 5 areas and usual census methods for other residents. However, after reviewing initial data files, the Census Bureau opted in 2019 not to rely on records in these areas—dissolving TEA 5 and allocating its BCUs to TEA 1 (Self-Response) and TEA 6 (Update Leave) as appropriate.

Copyright National Academy of Sciences. All rights reserved.

C.5 CENSUS PARTNERSHIP, OUTREACH, AND PROMOTION ACTIVITIES

C.5.1 2020 Census Partnership Program

The goal of the 2020 Census Partnership Program was to motivate national, regional, and local organizations to assist the Census Bureau in urging people in their networks to respond to the census mailings or to cooperate with census field personnel. To this end, the associated Census Media Campaign ran local paid media in every market in the United States, in addition to national media, to request people to respond to census materials and personnel. Further, the Census Bureau relied on a number of programs to communicate the importance of cooperation with the 2020 Census attempts at enumeration, described below.

Community Partnership and Engagement Program

The aim of the Community Partnership and Engagement Program (CPEP) was to enroll community partners to increase participation by people less likely to respond and often missed. The goals were to: (1) educate people about the 2020 Census and foster cooperation with enumerators; (2) encourage community partners to motivate people to self-respond; and (3) engage grassroots organizations to reach out to hard-to-count groups and those not motivated to respond. The CPEP partner organizations formed 8,470 Complete Count Committees, hosted more than 493,000 events, and completed more than 732,000 commitments to support the 2020 Census. CPEP partners and National Partnership Program partners (see below) organized activities such as digital weekends of action, recruiting events, Mobile Questionnaire Assistance events, Get Out the Count activities like block parties and caravans, phone banks, and text messaging campaigns. The partners also invited Census Bureau personnel to speak at their events. Further, the Census Bureau created more than 250 types of materials—fact sheets, posters, and social media toolkits—which were posted on the 2020 Census website and available in multiple languages. In addition, the Census Bureau shared information, resources, and operational updates with partners via an email list with nearly 80,000 subscribers. Local Partnership Specialists (Census Bureau staff hired specifically to facilitate CPEP) educated partners and stakeholders on updates and supported activities that encouraged households to respond.

National Partnership Program

The National Partnership Program (NPP) had very similar goals as the CPEP, but it was targeted toward national organizations rather than local organizations. The NPP signed up 1,090 national organizations (exceeding its goal of 900 organizations) by April 20, 2020.

Copyright National Academy of Sciences. All rights reserved.

Statistics in Schools

Regional Partnership Specialists partnered with local school districts, individual schools, Head Start programs, daycare centers, diaper pantries, and other community organizations to promote Statistics in Schools and the Counting Young Children initiative.

C.5.2 Integrated Partnership and Communications Operation

The goal of the Integrated Partnership and Communications Operation (IPC) was to communicate the importance of participating in the 2020 Census to the entire population of the United States, Puerto Rico, and the Island Areas. The hope was to engage and motivate people to self-respond, preferably using the internet. Also, during later stages of the census, the goal was to raise and maintain awareness of the census throughout the full 2020 Census process to encourage response and cooperation. The IPC Operation ran local paid media in every market in the United States in addition to paying for national media. The IPC Operation had four phases: Strategic Early Education Phase, Awareness Phase, Motivation Phase, and Reminder/NRFU Phase.

- *Strategic Early Education Phase* (January–December 2019): The goal of this phase was to build public trust in Census Bureau efforts to encourage participation, and to educate residents about the importance of accurate counts. The strategy was to utilize the CPEP, Statistics in Schools, and various public relations outreach to get this message across.
- *Awareness Phase* (planned for January 14–March 12, 2020; occurred January–February 2020): The goal of this phase was to notify the public about the upcoming 2020 Census, the purpose of the census, its importance, and how to respond to it. The strategy was to provide information on available means for completing the 2020 Census, explain where residents could access additional information and resources, and explain what residents could expect from the Census Bureau and its partners.
- *Motivation Phase* (planned for February–April 2020; occurred March 13–August 2, 2020, due to COVID-19-related adjustments): The goal of this phase was to work toward census completion by informing residents that the 2020 Census was underway and reminding them to participate in one of several ways. The strategy was to deliver general and audience-specific messages to encourage participation, as well as to urge people to encourage friends and family members to participate. In particular, the hope was to focus on "fence-sitters" who were comfortable responding via the internet but may not yet have decided to do so.
- *Reminder/NRFU Phase* (planned for May–August 2020; occurred August 3–September 27, 2020): The goal of this phase was to remind residents

Copyright National Academy of Sciences. All rights reserved.

that the 2020 Census was taking place and encourage them to participate if they had not yet done so. The strategy employed was dependent on the availability of response modes.

C.5.3 Other Media and Outreach Efforts

Targeted E-mail Outreach

In this operation during July–August 2020, 48 million targeted e-mails were delivered to low-responding regions, resulting in 7.8 million digital impressions of motivational videos.

2020 Census Barriers, Attitudes, and Motivators Study

In the lead-up to the 2020 Census, the Census Bureau initiated an effort to better understand attitudes and behaviors that underly various levels of census participation across demographic and geographic groups. The Census Barriers, Attitudes, and Motivators Study (CBAMS) consisted of a population survey of 50,000 households (35% response rate) fielded in 2018 and 42 focus groups in 14 cities. The general findings of this study were that motivators of census participation include the prospect of proper funding for community needs and services, such as additional hospitals, fire departments, and schools. CBAMS discovered a general lack of knowledge about the decennial census' scope, purpose, and constitutional foundation. Further, there was a general apathy towards the census. There were also concerns due to privacy worries and distrust of government. These knowledge gaps had important subgroup differences.

Changes Due to COVID-19

The emergence of COVID-19 during the 2020 Census created problems for messaging and outreach efforts. First, the media campaign had to be shifted to account for the adjusted messaging and deadlines. Second, the pandemic made some of the outreach efforts more difficult to arrange. Third, the pandemic dominated various media communications, so it was difficult to prioritize intended census messaging.

C.6 SELF-RESPONSE AND OTHER TYPES OF ENUMERATION OPERATIONS

In most areas of the country, the 2020 Census, like previous censuses, strove to motivate self-response by households through a series of mailings (in addition to the outreach and media campaigns described in Section C.5). In some areas and for some types of living situations, census field staff played a

Copyright National Academy of Sciences. All rights reserved.

Table C.1 Modification of Timing of 2020 Census Enumeration Operations Due to COVID-19

Operation	Original Timing	Actual Timing
Remote Alaska	1/21–4/30	1/21–3/18; 4/13–8/28
Self-Response Phase	3/12–7/31	3/12–10/15
Update Leave	3/15–4/17	5/4, 6/12–8/10
Update Enumerate	3/16–4/30	6/14–8/31
Service-Based Enumeration	3/30–4/1	9/22–9/24
Targeted Non-Sheltered Outdoor Locations	3/31–4/1	9/23–9/24
Group Quarters Enumeration (Residential)	4/2–6/5	4/2–9/3
Enumeration of Transitory Locations	4/9–5/4	9/3–9/28
Nonresponse Followup	5/13–7/31	7/16, 8/9–10/15

SOURCE: Fontenot (2021h:Slide 13).

major role in the enumeration: rural enumeration strategies (Update Leave, Update Enumerate, Remote Alaska); and GQ enumeration strategies (including Service-Based Enumeration and targeted non-sheltered outdoor locations, as well as the related Enumeration at Transitory Locations). Field staff also conducted Nonresponse Followup (NRFU)—see Section B.7. COVID-19 affected all these operations, to a greater or lesser degree, as summarized in Table C.1. The delays affected subsequent operations, including data processing and release.

C.6.1 Household Contact Strategies for Self-Response

2010 Census

A major difference between the 2010 and 2020 Censuses was that, in 2010, paper questionnaires were the primary medium used for collecting data. Questionnaires were mailed to U.S. households in the Mailout-Mailback TEA using a mailing strategy consisting of multiple contacts, as follows. First, an advance letter was sent, informing households that the census questionnaire would be arriving at their households soon. Then the initial questionnaire was sent, followed by a direct mail postcard sent to select areas. Then an English-only replacement questionnaire was distributed using one of three treatments: (1) blanket replacement mailing; (2) targeted replacement mailing; and (3) no replacement mailing. During the 2010–2020 intercensal period, the Census Bureau tested a variety of alternative strategies to improve the procedures used in the 2010 Census, drawing in particular upon the "wave methodology" used to promote self-response in recent Canadian censuses.

Copyright National Academy of Sciences. All rights reserved.

2020 Census

The goal of the Internet Self-Response (ISR) operation used in the 2020 Census was to communicate to the U.S. population the importance of self-responding, with the hope of generating the largest possible frequency of self-response. This would save considerable funds by reducing the need to conduct expensive in-person follow-up. Towards this end, the Census Bureau developed communication and contact strategies to encourage the use of the internet as the primary response mode, through a sequence of invitations and reminder mailings. Further, the goal was to increase the percentage of high-quality responses by increasing the opportunity to respond and the flexibility of response-mode options.

The first part of the mail strategy was to divide TEA 1 areas into "Internet First" and "Internet Choice" areas. Internet First areas were those for which the majority of residences had relatively easy access to the internet—specifically, census tracts not characterized as Internet Choice (see below). For those Internet First areas, the plan was to start with a letter from the Census Bureau, followed by a second letter, then a postcard, followed by a letter with a questionnaire included, and lastly an "It's not too late" postcard. Each letter informed the residents about how to be enumerated online.

Internet Choice areas were census tracts with low overall self-response rates in the ACS and one of the following: low ACS internet response rates, high rates of people ages 65 and older, or low rates of internet subscribership, all determined using 5-year ACS data for 2013–2017. The sequence of mailings for Internet Choice areas included an initial letter with a questionnaire enclosed plus a link to respond online, followed by a letter, then a postcard, followed by a letter with a questionnaire, and then an "It's not too late" postcard. Lastly, nonrespondents were mailed a final reminder postcard.

The timing of the originally planned contact strategy had to be modified as a result of the COVID-19 pandemic. Table C.2 provides the initial timings and adjustments made due to the pandemic.

In response to the pandemic, the Census Bureau used a sixth mailing (addressed postcards including a Census ID for internet response) sent to an estimated 50 million nonresponding housing units as of mid-July. There was also a seventh mailing with a paper questionnaire that was sent to about 16.2 million households in low-responding census tracts, targeted to households that had received only one paper questionnaire in the mail. In addition to the additional sixth and seventh mailings indicated above, there was an "Every Door Direct Mail" postcard, which was sent to 4.5 million USPS boxes in areas where the physical addresses of housing units were not eligible for any form of USPS carrier delivery service.

Also in response to COVID-19, the Census Bureau had to adjust 2020 Census operations to protect the health of respondents and Census Bureau field

Copyright National Academy of Sciences. All rights reserved.

Table C.2 2020 Census Self-Response Mail Contact Strategy and Adjustments for COVID-19

Internet First	(1) Letter	(2) Letter	(3)[a] Postcard	(4)[a] Letter Plus Questionnaire	(5)[a] Postcard	(6)[b] Postcard	(7)[c] Letter Plus Questionnaire
Cohort 1	3/12	3/16	3/26	4/8 (4/14)	4/20 (4/27)	7/22–7/28	8/21–9/14
Cohort 2	3/13	3/17	3/27	4/9 (4/18)	4/20 (4/30)	7/22–7/28	8/21–9/14
Cohort 3	3/19	3/23	4/2	4/15 (4/22)	4/27 (5/4)	7/22–7/28	8/21–9/14
Cohort 4	3/20	3/24	4/3	4/16 (4/24)	4/27 (5/6)	7/22–7/28	8/21–9/14

Internet Choice	(1) Letter Plus Questionnaire	(2) Letter	(3)[a] Postcard	(4)[a] Letter Plus Questionnaire	(5)[a] Postcard	(6)[b] Postcard
Cohort	3/13	3/17	3/27	4/9 (4/28)	4/20 (5/9)	7/22–7/28

[a] Targeted only to nonresponding households.
[b] Common postcard targeted to nonresponding housing units as of mid-July, including Census ID for internet response.
[c] Targeted to nonresponding housing units in low-responding census tracts that had received only one paper questionnaire in the mail.

NOTE: Dates are start dates for each contact; dates in parentheses are rescheduled start dates due to COVID-19. Only mailings (1)–(5) were part of the original planned contact strategy.

SOURCE: Fontenot (2021h:22–23); Hearns (2021:9–10); Fontenot (2021e).

Copyright National Academy of Sciences. All rights reserved.

personnel. New procedures had to be consistent with COVID-19 guidance from federal, state, and local authorities, without compromising the need to ensure a complete and accurate count. The Census Bureau monitored the rapidly changing conditions at the state and local levels and, in consultation with appropriate officials, updated the start dates for selected operations.

2020 Internet Instrument Design to Encourage Self-Response

Internet respondents received the following message on the 2020 Census online response landing page:

Welcome to the 2020 Census

- It's quick and easy. The 2020 Census questionnaire will take about 10 minutes to complete.
- It's safe, secure, and confidential. Your information and privacy are protected.
- Your response helps to direct billions of dollars in federal funds to local communities for schools, roads, and other public services.
- Results from the 2020 Census will be used to determine the number of seats each state has in Congress and your political representation at all levels of government.

Getting started:

- You must complete your questionnaire once you begin. If you leave the questionnaire and return later, you will have to start over.
- Do not use the web browser buttons (back, forward, of close browser). Use the buttons within the questionnaire to navigate.
- For best results, use the latest version of Chrome, Firefox, Internet Explorer, or Safari. Enable cookies.

The login page asked for a 12-digit Census ID which could be found in the delivered census literature, but the ID was not required for response. People responding without the Census ID (Non-ID cases) were directed to a page on which they entered street address information for the place where they usually resided on April 1, 2020, before proceeding with the rest of the questionnaire. The address information in these Non-ID cases would be matched to the MAF to mark the address as complete or to designate it for further resolution.

2020 Census Summary of Self-Response

The original dates for Self-Response were March 12–July 31, 2020, with July 31 being the original target end date for NRFU as well. The COVID-19 pandemic resulted in several shifts in the schedule for 2020 Census data collection, with Self-Response and NRFU both ultimately closing on October 15, 2020. Addresses that were ultimately enumerated through self-response comprised 67.0% of the total, which exceeded the 66.5% for those enumerated

Copyright National Academy of Sciences. All rights reserved.

in 2010.[5] Of those resolved in the 2020 Census, 79.7% were through online enumeration, 18.1% by paper, and 2.1% through use of the telephone.[6] While the operations in 2020 and 2010 were not directly comparable, in 2020 all 50 states and the District of Columbia met or exceeded the percentage of addresses resolved by Self-Response in 2010. By mode, self-response rates, adding to 67.0%, were as follows: internet, 53.5%; paper, 12.3%; and phone, 1.2%. Finally, 93 million responses were collected via ISR of which 76 million had Census IDs and 17.6 million were without Census IDs.

Non-ID Processing

The Non-ID Processing operation enabled households to potentially be enumerated via Self-Response without having a Census ID (hence the designation of Non-ID). The goal was to make it easier and more convenient for households to respond to the census via the internet even when they had not received or had misplaced the various mailings from the Census Bureau, which included an identification code or Census ID. The goal was to match respondent-provided Non-ID addresses to the census living quarters address inventory, along with attempts to assign any remaining nonmatching addresses to census geographic units.

Some form of Non-ID Processing was used in the 2010 Census, through the availability of "Be Counted" forms in public locations, but solely on paper. The use of the internet as primary response mode for the 2020 Census caused a major expansion in the number of Non-ID cases, from 2.9 million in 2010 to 17.6 million in 2020. Non-ID Processing was central to the promotion of the 2020 Census as being completable by respondents at any time and from any device. The switch to internet collection also allowed for some improvements in Non-ID work. First, the internet-response instrument allowed for consistent address collection during ISR and Census Questionnaire Assistance (telephone) interviewing, facilitating immediate edits and quality checks on provided addresses. Second, real-time address matching during internet and telephone response supported a feedback loop in which a respondent could potentially produce a better address to increase the potential for matching. Third, the use of administrative records data could sometimes correct erroneous address information provided by a respondent (e.g., a misspelled street name) or supply omitted data (e.g., an apartment number).

Table C.3 shows the disposition of Non-ID Processing cases—whether the respondent-supplied addresses matched to the MAF or not—in the 2020 Census.

[5]The denominator for these percentages is total addresses in the 2020 MAF, including addresses that turned out to be occupied or vacant housing units and addresses that turned out to be nonresidential or nonexistent.

[6]Many of those who called Telephone Questionnaire Assistance were asked during the call whether they wanted to be enumerated on the call.

Copyright National Academy of Sciences. All rights reserved.

Table C.3 2020 Census Non-ID Workload by Process and Outcome

Outcome	Percent
Total Non-ID Workload—17.6 million households	100.0
Matched	91.2
Matched via Automation	83.9
Matched in Real Time	83.3
Matched Asynchronously	0.6
Matched Clerically	7.3
Nonmatched, geocoded, and verified	5.2
Nonmatched, unable to verify	3.6

SOURCE: Donello (2021:Slide 13).

The 5.2% of the total 17.6 million addresses that could not be matched to the MAF but that were ultimately able to be geocoded and verified is central, as it represented an important set of previously unknown addresses to be added to the MAF.

Coverage Improvement Program

The 2020 Census Coverage Improvement Program was an attempt to resolve possible erroneous enumerations and omissions, including people missed, counted at the wrong place, counted in error, or counted more than once, through contacting households that had responded but may have had coverage errors. The Census Bureau used respondent-provided telephone numbers to conduct telephone follow-up interviews in specific kinds of households to determine whether the household rosters on their census returns should be updated. Respondents were probed to identify whether residents were missed or counted in error. The hope was to identify people who were not living or staying at the address on Census Day (April 1, 2020), or who were missing from the original roster but living or staying at the address on Census Day. There was also an attempt to collect address information for people who indicated that they usually lived or stayed somewhere else.

Households were identified for follow-up through various checks. One possibility was a difference between the number of names on the roster and the population count of the household. Second, if the undercount question was answered affirmatively, then the household was followed up. Third, if someone usually lived elsewhere (e.g., in a type of GQ, at a seasonal home, with relatives, or for a job), the household was followed up.

Copyright National Academy of Sciences. All rights reserved.

C.6.2 2020 Census Rural Enumeration Operations: Update Leave, Update Enumerate, and Remote Alaska

2020 Update Leave

Update Leave (UL) occurred in geographic areas (TEA 6) characterized by one or more of the following: (1) housing units did not have city-style addresses; (2) most people received mail at post office boxes or at drop points; (3) the area had been affected by natural disasters; or (4) the area had high concentrations of seasonally vacant housing. The objectives of the 2020 UL program were to update the address and feature data for each housing unit and to leave an Internet Choice questionnaire package at every housing unit beginning on March 15, 2020. UL personnel used the LiMA software and device that was used for 2020 IFAC. The questionnaire packet that was used in areas with city-style addresses was expanded for this purpose to include all modes of Self-Response: online, phone, or questionnaire to be mailed back. Questionnaires contained a generic ID and barcode that represented the housing unit's response once linked to the address during the update process. Two reminder mailings were planned: (1) a letter mailed on or around April 1 with the associated ID for the address to respond online; and (2) a postcard, mailed on or around July 17, to any households from the first mailing that did not respond. Housing units in UL areas that did not self-respond were sent to the NRFU operation for follow-up.

UL operations were paused on March 18, 2020, due to COVID-19, at which point work was completed for only 10.8% of the addresses and quality control work had not yet begun. In coordination with state and local health officials, a phased restart began the week of May 4 with the reminder mailings rescheduled accordingly. Operations initially resumed in 23 Area Census Offices (ACOs) in 13 states. All ACOs resumed UL operations by the week of June 11. By July 16, nearly all (99.5%) of UL addresses were enumerated or otherwise resolved, and the remainder were completed by August 10. UL resulted in paper questionnaires and invitations to respond online being delivered to over 6.8 million addresses.

2020 Update Enumerate

The Update Enumerate (UE) operation was initially planned to occur in areas (TEA 2) where the initial visit required enumeration while updating the address frame. The primary functions of UE included: (1) verifying and updating the address list and feature data; (2) determining the type and address characteristics for each living quarter; and (3) enumerating respondents at housing units. UE occurred in areas that were part of the 2010 Census Remote Update/Enumerate operation such as northern parts of Maine and southeast Alaska, along with select tribal areas that requested to be enumerated in person during the initial visit. If there was no contact made with the residents, the field

Copyright National Academy of Sciences. All rights reserved.

staff left a notice of visit that was linked to the address for the housing unit to respond online.

Original plans for the 2020 Census had UE as the primary TEA alternative to Self-Response. However, in May 2017, the Census Bureau decided to reinstate the UL program to the 2020 Census, with most of the addresses planned for UE being switched to UL. The original operational dates for UE were March 16–April 30, 2020, but due to the COVID-19 pandemic, UE was suspended on March 18 and only resumed on June 14. In the revised UE operation, field staff used masks, gloves, and hand sanitizer, and maintained a distance of at least 6 feet during enumeration. The UE operation concluded on August 31.

2020 Remote Alaska

The objectives of the methodology used in the Remote Alaska operation (TEA 4) were the same as those for the 2020 UE operation. This operation focused on address updating and enumeration of Alaska Native villages, and it included advance contact and enumeration of the GQs and TLs within the Remote Alaska area. The initial workload was over 28,000 living quarters. The original operational dates were January 21–April 30, 2020. Due to the COVID-19 pandemic, the operation was halted on March 18, resumed on April 13, and finished on August 28.

C.6.3 Enumeration of Group Quarters in the 2020 Census

A GQ is a place where people live in a group living arrangement that is owned or managed by an entity. The services provided may include custodial or medical care, and residency is often restricted to people receiving those services. People living in GQs are usually not related to each other. GQs include such places as college/university student housing, residential treatment centers, skilled nursing facilities, group homes, correctional facilities, military barracks or ships, and workers' group living quarters.

As part of the GQ operation, the Census Bureau developed special enumeration procedures to provide opportunities for people experiencing homelessness to be counted at service locations and pre-identified outdoor locations. These were soup kitchens, regularly scheduled mobile food vans, and targeted non-sheltered outdoor locations. In all, there were close to 30 types of GQ codes and definitions.

The GQ population comprised about 2.5% of the total population in 2020. While this is a relatively small percentage overall, GQs can comprise a substantial percentage of the jurisdictions in which they are located (especially for correctional institutions and college and university populations).

For all GQs, the Census Bureau endeavored to identify a GQ administrator, in some cases working through national associations and organizations to

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

obtain contact information for administrators of the various GQ types. The administrator was to help the Census Bureau decide the type of GQ, how to gain access to residents, and how the residents would prefer to be counted. As an example, assisted living facilities are considered housing units, but any floor or wing where residents receive 24 hours of care 7 days per week (e.g., skilled nursing) are themselves considered GQs. As a result, some of the residents would have received an invitation to respond via the internet or mail, while the administrator would decide how the GQ residents would participate in the census.

Some student housing facilities are considered housing units, while some are considered GQs. If a hotel is considered a person's "usual home elsewhere" (the place other than the GQ/TL where the person lives and sleeps most of the time), it is considered a housing unit, but it is a GQ if the hotel houses people because of a natural disaster. If all units of the hotel are housing individuals experiencing homelessness, it is considered a GQ. If a hotel is housing both individuals experiencing homelessness and individuals who are paying for extended care, it is considered a housing unit.

Strategies for Enumeration of Residential Group Quarters

Residential GQs represented a challenge to census enumeration. There were different enumeration strategies for many types of residential GQs. Also, GQ administrators for several types of GQs had more than one option that could be used to count residents, including various paper data-collection methods and an electronic data-collection method. Strategies were designed to be flexible because GQ administrators sometimes needed to change enumeration methods at the time of enumeration.

GQ enumeration procedures used in the 2020 Census included: (1) in-person interviews; (2) drop-off/pick-up of questionnaires; (3) facility self-enumeration; (4) paper listings sent to enumerators to send to ACOs; (5) mail-out/mail-back of questionnaires from ACOs to GQ administrators and back; and (6) eResponse—the electronic transfer of data records, which entailed GQ administrators uploading administrative response data for individuals living or staying in their facility to a secured web-based Census Bureau portal.

The three key principles of the criteria for GQ enumeration in 2020 were essentially unchanged from 2010: (1) count people at their usual residence—the place where they live and sleep most of the time; (2) count people in certain types of group facilities on Census Day at the group facility; and (3) count people who do not have a usual residence, or who cannot determine a usual residence, where they are on Census Day. The 2020 Census GQ operations included: (1) Group Quarters Advance Contact (In-Office/In-Field) and Enumeration; (2) Service-Based Enumeration; (3) Maritime/Military Vessel Enumeration; and (4) Deployed Military Enumeration.

Copyright National Academy of Sciences. All rights reserved.

Group Quarters Advance Contact and Enumeration

Group Quarters Advance Contact ran from February 3–March 6, 2020. Area Census Office clerks called GQ administrators to confirm and explain the procedures to be used for the upcoming GQ enumeration. Census workers then verified the GQ's name, address information, contact name, phone number, and business email address. For the upcoming Census Day, census workers collected an expected population count and maximum population count. Census workers also helped the facility manager to decide on the method of enumeration based on the GQ type and scheduled an appointment for enumeration. If there was no available contact name or phone number, census staff made a personal visit to collect the above information and to discuss the enumeration procedures to be used.

GQ enumeration was planned to run from April 1–June 5, 2020, and eResponse data collection was planned to run from April 1–May 1, 2020. Due to COVID-19, the dates for GQ enumeration were extended to August 26, 2020, as was eResponse and mailback of questionnaires.

COVID-19 complicated census enumeration in the GQ population. Various GQs (e.g., nursing homes, hospitals, other health-based facilities, colleges/universities, group homes) had restrictions on access, some colleges/universities were closed, some enumerators refused to work, individuals in quarantine on military bases or ships were more difficult to account for, and staff were sometimes not available to handle specific planned activities.

For colleges and universities, GQ administrators originally selected methods that allowed students to self-respond as in previous censuses. In terms of the approaches that were initially agreed upon:

- 47% chose the eResponse methodology (administrators would upload directory-level information about each student onto the Census Bureau's secure server);
- 35% chose the drop-off/pick-up option (census field staff would drop off individual census questionnaires for the administrator to distribute to students, after which census field staff would return to pick up the completed questionnaires);
- 9% chose in-person interviews (census field staff would conduct interviews with students to complete the questionnaires); and
- 10% chose paper response data collection (administrators would provide census field staff with a paper listing containing directory information about their students).

However, many students were sent home due to COVID-19 and were not living in student housing at the time of GQ enumeration, so the Census Bureau staff reached out to GQ administrators that chose the in-person and drop-

Copyright National Academy of Sciences. All rights reserved.

off/pick-up methods to ask them to choose either eResponse or the GQ paper-response data-collection option. Based on feedback from the U.S. Department of Education, the eResponse and paper-response methods had implications for the types of data that could be provided under the Family Educational Rights and Privacy Act of 1974 (FERPA). Schools could not provide "Directory Information" for students without their consent. Also, schools could not include race, ethnicity, or gender as directory information under FERPA. The Census Bureau believes it counted 88% of college students during the regular GQ enumeration data collection. Post data collection (December 2020), the Census Bureau telephoned student housing facilities and other GQ types to verify counts and attempt to convert refusals. As a result, college/university student housing enumeration increased from 88% to 97%.

Service-Based Enumeration, Targeted Non-Sheltered Outdoor Locations Enumeration, and Enumeration at Transitory Locations

Enumeration was conducted at various places such as soup kitchens and homeless shelters and targeted non-sheltered outdoor locations (TNSOLs) to enable people experiencing homelessness to be enumerated in the census. Locations where people might be staying on a transitory basis were also visited in the Enumeration at Transitory Locations (ETL) operation to determine their usual residence (elsewhere) or to enumerate them at the location if they had no other usual residence. Service-Based Enumeration (SBE) was originally scheduled to be conducted March 30–April 1, 2020, TNSOL enumeration was originally scheduled to be conducted March 31–April 1, 2020, and ETL was originally scheduled to be conducted April 9–May 4, 2020.

In late May to early June and in response to the COVID-19 pandemic, Census Bureau representatives consulted with major stakeholders to determine alternative dates to conduct SBE/TNSOL. Determining optimal dates to conduct these activities considered the need to conduct a thorough and accurate enumeration, while also understanding the needs of external partners that were critical to conducting SBE. These discussions resulted in the decision to conduct SBE/TNSOL in fall 2020. A primary reason had to do with the effects of seasonality on the homeless population. Late September approximates the weather that would have been experienced in March. Summer would be a more difficult time to enumerate the homeless population since fewer people use shelters and people experiencing homelessness are more spread out. In addition, at the time, it was anticipated that COVID-19 infections could recede by fall. Finally, the service providers at SBE sites were also contending with the effects of the COVID-19 pandemic, and waiting until fall gave them time to prepare. Over 53,000 service-based locations were ultimately enumerated, including almost 37,000 TNSOLs.

Copyright National Academy of Sciences. All rights reserved.

The rescheduled SBE/TNSOL operations were conducted September 22–24, 2020. The service-based locations included shelters for people experiencing homelessness, soup kitchens, and regularly scheduled mobile food vans. The primary enumeration methods were in-person interviews and paper listings, in which field staff picked up paper listings from GQ administrators. Drop-off/pick-up of questionnaires was an added option for shelters only. The primary method of enumeration for ETL was in-person interviews, and this operation was completed between September 3–28, 2020.

Maritime/Military Vessel and Deployed Military Enumeration

The Maritime/Military Vessel Enumeration was planned to run from April 1–June 20, 2020. Due to COVID-19, the end date for Maritime/Military Vessel Enumeration was extended to August 24, 2020. Military deployed overseas temporarily and military stationed overseas were enumerated using administrative records.[7]

C.7 NONRESPONSE FOLLOWUP

In TEAs 1 and 6, NRFU was a critical part of census operations. The goal of the NRFU operation was to enumerate people living at an address for which no self-response had been received (or a determination that an address was vacant housing or should be deleted from the census because it no longer existed or was nonresidential or uninhabitable). NRFU was originally planned to run from May 13–July 31, 2020. The advent of COVID-19 necessitated that these dates be modified. The Census Bureau began a soft launch of NRFU on July 16, went to full implementation on August 11, and completed NRFU on October 15, 2020. This section describes the NRFU contact strategy, identification of vacant and nonexistent addresses, outcomes of NRFU, additional impediments for NRFU in 2020, a related operation (census questionnaire assistance), and use of administrative records for enumeration.

C.7.1 NRFU Contact Strategy

The NRFU contact strategy had two primary purposes: (1) determine the housing unit status for nonresponding addresses; and (2) enumerate housing units for which a 2020 Census response was not received. Other activities

[7]Military temporarily deployed were included in the resident population for reapportionment, redistricting, and other data uses at their home base. Military stationed overseas were included in the population totals for their home states for reapportionment of the U.S. House of Representatives only and were not included in any other data product. See "The Census and the Military" overview slides at https://www.doi.gov/sites/doi.gov/files/uploads/oia-02052020-census-and-the-military.pdf.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

included: (1) field verification of addresses from respondents using Self-Response that did not contain a Census ID (Non-ID cases) and that were not on the MAF; (2) enumeration of UL cases that did not self-respond on first contact; and (3) resolution of additional cases from other census operations, such as LUCA appeals and the USPS DSF refresh.

Phase 1 of the NRFU strategy included all addresses that did not self-respond. All cases were assigned to an enumerator based on the optimization of work availability and workload geography. Use of proxies was approved on the third visit and continued through all following visits. If administrative records data could be found that were considered of sufficient quality, administrative records modeling was used to classify cases as occupied, vacant, or delete after one field visit. Up to four visits were made during Phase 1. All unresolved cases were put on hold after the fourth attempt in preparation for Phase 2.

Phase 2 was referred to as Semi-Permanent Assignment. Census Field Supervisor (CFS) Areas (the geography for which a single census field supervisor was responsible) were eligible to move into Phase 2 when, initially, 60% of cases had reached four visits or were completed. That percentage was raised to 85% on August 19, 2020, to ensure that CFS areas did not move into Phase 2 too soon. All CFS areas were eligible to move into Phase 2 on September 4, 2020, regardless of their completion percentage.

Census Field Managers used the Field Operations Control System to put CFS areas (approximately 4,500 housing units in each area) into Phase 2 and designated high-performing enumerators to work those areas. Reinterviews and self-response quality-assurance cases were prioritized in both Phases 1 and 2.

Phase 3 was referred to as Closeout. CFS areas were eligible to move into Phase 3 when a sufficient percentage of cases had either been resolved or had the maximum number of attempts at enumeration. Prior to the start of the operation, this threshold was set at 85%, but it was updated to 90% on August 19, 2020, to ensure that CFS areas did not move into Phase 3 too soon.

All CFS areas were eligible for Closeout on September 11, 2020, regardless of their completion percentage. During Phase 3, an additional round of administrative records modeling was used to reduce the NRFU Closeout workload. Cases modeled through use of administrative records as occupied, vacant, or delete were closed if they had six visits. Cases continued to be assigned using the same methodology used in Phase 2. CFS areas closed out upon 100% completion or the arrival of the last day of the operation on October 15, 2020.

C.7.2 Identification of Vacant and Nonexistent Addresses

In May 2020, administrative records were used to determine possible vacant and nonexistent addresses. To be determined to be vacant or nonexistent, an address had to have at least one "unavailable as addressed" (UAA) designation

Copyright National Academy of Sciences. All rights reserved.

from the USPS in mailable areas. If the mail was undelivered and there was no sign of occupancy, the address received one field visit. Otherwise, it received up to six field visits in NRFU.

The 2020 Census was the first to make use of administrative records and third-party data to assist in NRFU by identifying vacant or nonresidential units, and by using high-quality administrative records data to enumerate occupied households. Also, this was the first census that used an adaptive design strategy to reduce the number of contact attempts for each household.

C.7.3 Nonresponse Followup: Outcomes and Additional Impediments

Among housing units that were identified as occupied during NRFU, 55.5% were enumerated through an interview with a household member and 26.1% were resolved through interview of a proxy respondent. This is comparable to the 2010 Census, in which 24.7% of occupied housing units in NRFU were enumerated via proxy. Finally, 18.4% of the occupied housing units in NRFU were enumerated using high-quality administrative records.

For housing units that were determined to be vacant in NRFU, 85.5% were resolved through an interview with a proxy respondent and 14.5% were resolved through use of administrative records. Among those housing units that were determined to be deleted (nonexistent or nonresidential) in NRFU, 96.4% were resolved with a proxy respondent and 3.6% were resolved using administrative records.

Every decennial census encounters circumstances that make NRFU and other operations more difficult than anticipated, but the delay in starting 2020 field work raised particular concerns for NRFU—putting the most intricate of census operations in the midst of peak natural disaster season, which the regular schedule seeks to avoid. In 2020, the circumstances facing NRFU included (1) the emergence of COVID-19 in March 2020—accompanied by various local rules and regulations regarding public behavior, in-person contact, and availability of personal protective equipment; (2) Tropical Storm Marco, which hit the Gulf Coast on August 24, 2020; (3) Hurricane Laura, which hit the Gulf Coast on August 26, 2020; (4) California, Oregon, and Washington wildfires and associated poor air quality in early September 2020; (5) Hurricane Sally, which hit the Gulf Coast on September 16, 2020; (6) Tropical Storm Beta, which hit the Gulf Coast on September 21, 2020; and (7) various legal challenges to the timing of various census field operations.

C.7.4 Census Questionnaire Assistance and Telephone Response

The two goals of the 2020 Census Questionnaire Assistance (CQA) operation were: (1) provide questionnaire assistance for respondents by answering questions about specific items on the census form or other frequently

Copyright National Academy of Sciences. All rights reserved.

asked questions about the census; and (2) provide an option for respondents who were interested in conducting their census interview over the telephone. A similar program was implemented during the 2010 Census, however, there were substantial changes in implementation for 2020. For example, during the 2010 Census, responses were collected primarily from paper questionnaires, and telephone support centered around supporting the paper operation; whereas the primary mode for Self-Response in 2020 was ISR, so CQA was intended to meet the needs of respondents who did not have access to the internet or computers. There were other typical questions anticipated about the 2020 questionnaire in 2020 in terms of how to respond, so toll-free telephone numbers were made available for respondents to call for help in completing their responses, which they could do over the telephone.

Other types of assistance for respondents were interactive voice response (IVR) and Frequently Asked Questions. Language support was also provided in English, Spanish, Tagalog, Arabic, Haitian Creole, Polish, French, Chinese (Mandarin and Cantonese), Korean, Russian, Vietnamese, Japanese, and Portuguese. English and Spanish were the predominant languages that were used to provide assistance. CQA received 13.5 million calls. Of these calls, 87.4% were received on English language lines and 10.4% on Spanish language lines. A total of 12.9 million calls accessed the IVR, and 7.9 million calls were handled within the IVR.

C.7.5 Use of Administrative Records for Enumeration

A major initiative of the 2020 Census intercensal research and development program was to make possible the use of administrative records in enumeration. Uses included: (1) administrative records modeling to support the reduction of contacts in the NRFU operation; (2) the assignment of a Census ID to self-respondents who did not have their census ID; and (3) Self-Response and NRFU quality assurance. This section focuses on the first application (reducing contacts during NRFU) and on how administrative records modeling was modified based on the delayed start of the NRFU operation due to COVID-19 and the extension of the Internal Revenue Service (IRS) tax filing deadline from April 15 to July 15.

Using Administrative Records to Identify Vacant and Nonexistent Addresses

Part of the original 2020 Census plan was to use administrative records data to help identify vacant and nonexistent addresses and, as a result, minimize field work to validate those determinations. Sources of administrative records included: USPS designations of UAA for census mailings made around April 1, 2020; USPS DSF and other USPS information; IRS 1040 filings and 1099

Copyright National Academy of Sciences. All rights reserved.

information returns; the Centers for Medicare and Medicaid Services (CMS) Medicare Enrollment database; the Indian Health Service Patient database; the Veterans Service Group of Illinois third-party national files; the Census Bureau's MAF; and ACS area-level estimates of the percent of addresses that were vacant, in poverty, and occupied by people with other characteristics. The original 2020 Census plan was to use these sources to help identify vacant and nonexistent addresses. The Census Bureau developed a distance function, which measured the degree to which information from the above sources discriminated a vacant address from a nonexistent address and both from an occupied address. The resulting algorithm found that, to be considered for possible vacant or nonexistent status, an address had to have at least one UAA in the census mailings.

The original plan was to send a mailing to each nonresponding address that the administrative records model determined to be possibly vacant or nonexistent, about six weeks after Census Day (mid-May). This was to be followed up by a single field visit. If the mail was undelivered and there was no sign of occupancy, the administrative records model was to be used to decide on the status of the address as vacant, nonexistent, or possibly occupied. If the mail was delivered or if there was a sign of occupancy, the address was placed into NRFU.

Due to COVID-19, there were changes to the tax filing deadline and to the NRFU deadlines. On March 21, 2020, IRS announced that the tax filing deadline was extended from April 15, 2020, to July 15, 2020. Throughout, however, the Census Bureau continued to receive monthly deliveries of processed IRS-1040 records. The start of the NRFU operation was also delayed until August 9, 2020. As a result, the Census Bureau revised the plan for use of administrative records modeling for possible vacant or nonexistent addresses. The Census Bureau delayed determination of vacant and nonexistent addresses until the end of May 2020; it also scaled back the addresses that would be enumerated via administrative records as vacant or nonexistent, to err on the side of caution. For example, administrative records modeling was not used for addresses in ZIP codes with 57% or higher rates of UAAs on the first mailing; instead all of these addresses were sent to NRFU.

Enumerating a Percentage of Nonresponse Followup Workload Using Administrative Records

As part of NRFU, the plan was to use administrative records to enumerate some nonresponding housing units that had high-quality administrative records information. The idea was to build a roster from the most recent administrative records sources, check that multiple sources indicated that the household resided at the address in question, determine how likely it was that the census was counting all the people rostered at the right location, and then decide whether

Copyright National Academy of Sciences. All rights reserved.

to use administrative records data for the address in the census. The sources used for rostering were tax year 2019 IRS 1040 tax returns, tax year 2019 IRS information returns, the CMS Medicare Enrollment database, the Indian Health Service Patient database, and the Census Bureau's Household Composition Key File. A wider range of sources was used to verify that a household lived at a specific address.

The process for using administrative records information to enumerate households was planned to operate as follows. For TEA 1 areas, if there was no Self-Response return after the final (originally the fifth) mailing, the household would be visited by an enumerator and receive a postcard reminder one week later. If there was still no response, then the address would be enumerated using administrative records information, if that information was found to be of sufficiently high quality. For TEA 6 areas, an enumerator visited if there was no response from leaving off the questionnaire and after two reminder mailings.

With the delays in NRFU and tax filing, the Census Bureau shifted to June to identify initial administrative records-occupied cases in which there were also count agreements with Self-Response returns. The Census Bureau continued modeling, making one visit to an address, and enumerating households via administrative records if that information was deemed reliable; the 2020 Census also used tax return data delivered by IRS in June, July, and August.

Once a Census Field Supervisor area went into NRFU "Closeout," then the criteria for use of administrative records for enumeration of an occupied address were relaxed. A single source could be used that provided a population count, even if it lacked characteristics.

Finally, for students living in off-campus student housing, the Census Bureau decided in May to contact colleges and universities to see if they could provide names, dates of birth, and ages of students who were living off campus in the spring 2020 semester. Between June 18–August 14, the Census Bureau contacted over 1,300 schools to see if they could participate. As a result, over 600 schools provided their information to the Census Bureau.

C.8 DATA PROCESSING AND CREATION OF DATA PRODUCTS

C.8.1 2020 Census Data Files

The following is a list of internal files that represent steps in the processing of data collected in the 2020 Census leading up to the release of state population counts for reapportionment and block data for redistricting, including the dates when each file was originally planned to be completed and when it was actually created.

- **Decennial Response File 1 (DRF1):** The first file produced after data were collected was DRF1, containing all the response data, including duplicate responses. DRF1 is the complete inventory of every residential

Copyright National Academy of Sciences. All rights reserved.

address in the nation linked to every response received during data collection. Planned dates: 9/15/2020–10/15/2020; actual dates: 10/29/2020–12/26/2020.

- **Decennial Response File 2 (DRF2):** This file resulted from employing the Primary Selection Algorithm[8] to select which data from DRF1 should represent a housing unit on the Census Unedited File (CUF). The DRF2 universe was smaller than that for DRF1 because duplicate and multiple responses had been resolved. Planned dates: 10/16/2020–11/7/2020; actual dates: 12/26/2000–2/24/2021.
- **Census Unedited File (CUF):** After count imputation was applied to the remaining unresolved cases in DRF2, the result was the CUF. The CUF contained the final 2020 Census universes of housing unit records (including vacant units but not those deleted because they were nonresidential or nonexistent) and person records. The state-level population counts for apportionment are derived from the CUF. Planned dates: 11/8/2020–12/5/2020; actual dates: 2/25/2021–3/10/2021.
- **Census Edited File (CEF):** The CEF resulted when editing and item imputation were applied to missing and erroneous values for all items in the CUF records. Hence, the CEF contains complete data for all items. Planned dates: 11/25/2020–1/26/2021; actual dates: 4/20/2021–6/24/2021.
- **Microdata Detail File (MDF):** When confidentiality protection and recodes were applied to the CEF using the Disclosure Avoidance System (DAS), the result was the MDF. After tabulation geography was added, the data were ready for tabulation and dissemination. Planned dates: 1/26/2021–2/10/2021; actual dates: 6/25/2021–7/18/2021.
- **Tabulation:** Table-generation software transformed the MDF into tables for public release in the 2020 Census Redistricting File. Planned dates: 2/10/2021–3/19/2021; actual dates: 7/19/2021–8/12/2021.[9]
- **Apportionment and Redistricting Key Milestones:** The legislated deadlines are December 31 of the census year for apportionment counts and March 31 of the subsequent year for redistricting counts. Planned dates for apportionment preparation and release for the 2020 Census: 12/6/2020–12/27/2020; actual dates: 3/12/2021–4/26/2021. Planned dates for redistricting preparation and release: 2/17/2021–3/31/2021; actual dates: 8/13/2021–9/20/2021.

[8]The Primary Selection Algorithm is a decennial census algorithm that chooses among multiple enumerations for a given residence.

[9]The same process of creating an MDF and producing tables is being used for the Demographic and Housing Characteristics File (DHC); different processes, which skip the MDF stage, are being used for the Detailed DHC products (A and B) and the Supplemental-DHC product. See Chapter 11 for details.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

C.8.2 Count Imputation

The Census Bureau tries hard to reduce the frequency of households that fail to respond, but there always remain housing units known to be occupied but for which nothing about the household is known, including the number of residents. Further, there are housing units for which even the vacancy status is unknown. In these situations, the Census Bureau uses count imputation to estimate the number of residents on Census Day after checking administrative records for relevant information.

At the beginning of data collection, the objective is to assign a status (occupied, vacant, deleted) to every address in the MAF. Through Self-Response and NRFU, the Census Bureau strives to determine the count at every occupied address, which can result in two possibilities: (1) the address is *resolved*—field staff determine whether the housing unit is occupied (vacant, or delete, i.e., nonexistent or nonresidential), and if occupied, how many people live there; or (2) the address is *unresolved*—field staff have insufficient or conflicting information about the status of the address or how many residents live there. Unresolved addresses go to count imputation for resolution.

Types of Count Imputation

Count imputation may involve status imputation. If it is *unknown* whether a housing unit exists, is vacant, or is occupied, then count imputation can result in several possibilities: the housing unit is nonexistent, the housing unit exists but is vacant, or the housing unit is occupied with 1–9 or more people. If it is *known* that the unit exists, count imputation can result in the following possibilities: the housing unit is vacant, or it is occupied with 1–9 or more people. Finally, if it is *known that the housing unit is occupied*, count imputation can result in its being occupied with 1–9 or more people.

Count imputation is implemented after all data collection is complete (i.e., after Self-Response and NRFU). Data collection completion also includes the conclusion of administrative records enumeration since addresses with high-quality administrative records have already been resolved. Count imputation is a key step in creating the CUF.

Count imputation is a component of two important quality metrics. The first is the percentage of addresses that were unresolved at the end of data collection. The second is the percentage of the total resident population added through count imputation. These two percentages, when compared across decennial censuses by state and other geographies, are an indicator of comparative quality of census information.

Copyright National Academy of Sciences. All rights reserved.

2020 Count Imputation Procedure

In 2020, all addresses (resolved and unresolved) were placed into one of several mutually exclusive imputation cells, defined by characteristics related to operational paradata, such as: whether the address was a valid living quarters or not; whether it was likely a delete, vacant, or other; whether it was considered residential or not in the spring 2020 DSF; and whether it was UAA. Within each state and imputation cell, addresses were sorted by geography (state, county, census tract, block group, block).

For unresolved addresses, the count imputation procedure was conducted by copying the status, population count, or both from the values obtained from a resolved nearest neighbor housing unit that had the same values for the operational paradata of the imputation cell. By using the nearest neighbor, the count imputation procedure took advantage of spatial correlations among nearby addresses. A donor housing unit could be used as the nearest neighbor for more than one other housing unit.

A different procedure was used for count imputation for GQs, due to the differences and complexity of GQs in comparison with housing units. (Count imputation for GQs was an innovation in 2020 because of the problems COVID-19 caused for the GQ enumeration.) Several forms of review occurred to improve accuracy for GQs. These included calling GQs that had not provided a population count, calling GQs with larger-than-expected counts, a review of colleges/universities using data from the U.S. Department of Education, and a review to reduce duplicate people in GQs. If it was still not possible to obtain a count of residents, even though there was evidence that the GQ was open and occupied on Census Day (e.g., refusals and GQs that were closed during the visit for GQ enumeration), the Census Bureau used information such as the expected or maximum capacity reported for a GQ in the Advance Contact operations or the ACS to estimate occupancy on Census Day and, less often, information from administrative records or other GQs.

C.8.3 Item Nonresponse and Characteristic Imputation

Characteristic imputation is used in the census to eliminate item nonresponse. Because demographic information is used for tabulations, characteristic imputation is important for accuracy of cross-tabulations. The objective is for each person, housing unit, and GQ on the final CEF to have consistent and valid values for all items.

The census item nonresponse rate is essentially the proportion of missing responses before any pre-editing or imputation procedures are used for a given item. For example, if the respondent answers all questions for all household members except age, then age for each household member counts as an item nonresponse. If the respondent provides the number of people living at the

Copyright National Academy of Sciences. All rights reserved.

Table C.4 Item Imputation Rates (Percentages), 2010 and 2020 Censuses

Census/Item	Tenure	Sex	Age/DOB	Hispanic Origin	Race	Household Relationship
2010						
Imputed	3.5	1.6	5.1	4.5	4.1	2.1
Assigned	N.A.	1.3	1.5	1.7	1.2	0.5
Allocated	3.5	0.3	3.6	2.8	2.9	1.7
2020						
Imputed	7.6	6.2	14.8	8.7	9.2	8.0
Assigned	1.9	5.8	8.7	4.7	4.9	2.1
Allocated	5.7	0.4	6.1	4.0	4.3	5.9

SOURCE: DeJesus and Konya (2023:Table N).

housing unit but does not answer any of the demographic questions on the census questionnaire for any household member, all those characteristics count as item nonresponse. Inconsistent responses (e.g., responses incompatible with other responses) are not considered item nonresponse but are put through edit and imputation processes as needed.

2020 Edit and Characteristic Imputation Overview

In 2020, during the editing phase, responses were run through a series of checks to: (1) detect and correct out-of-range or inconsistent values; (2) remove invalid or duplicate responses (e.g., checking more than one relationship code on a paper questionnaire); (3) convert date of birth values to age values; and (4) assign race/Hispanic-origin responses to numeric codes. The optimal scenario was for all responses to be valid and consistent so there would be no need for imputation. When there were missing, invalid, or inconsistent values, the Census Bureau used characteristic imputation.

Characteristic imputation involves assignment and allocation. Assignment occurs when information can be determined from other responses provided for the same person, from previous census responses, or from administrative records for that person. Allocation occurs when responses that are missing or inconsistent are imputed using information from responses provided for other persons in the household or from a similar nearby household. Table C.4 provides item imputation rates for the 2010 and 2020 Censuses. Note that 2020 rates are higher than 2010 rates; enumerator returns in both censuses have higher imputation rates than do self-returns (data not shown).

Copyright National Academy of Sciences. All rights reserved.

Edit and Imputation Procedures for Specific Questionnaire Items

If a person's *sex* was missing, invalid, or inconsistent (e.g., if both male and female were checked), the Census Bureau attempted the following, in order: (1) use the first name for missing sex characteristics; (2) link the person to high-quality administrative records, such as their 2010 Census response or information from the Social Security Administration (SSA) (i.e., the Numident File); (3) assign sex to maintain household consistency; or (4) allocate sex from the nearest neighbor hot deck.[10]

If *age* or *date of birth (DOB)* were missing, invalid (e.g., over 150 years old), or inconsistent, there were several ways the Census Bureau could impute an age value. If either age or DOB was missing but not both, the Census Bureau calculated age from the DOB or created a random DOB from the age. If an inconsistent age and DOB were provided, the Census Bureau chose the response that was consistent with the person's relationship to the householder. Age and DOB could also be assigned from the person's responses to the 2010 Census or from the Numident File. If none of these options were available, the Census Bureau could allocate age and DOB from the nearest neighbor hot deck.

Some respondents did not provide Hispanic-origin or race information, or provided responses that had to be edited to a single Hispanic-origin response and up to eight race responses. The Census Bureau did not impute detailed Hispanic-origin and race responses when only a major group was reported. If Hispanic origin and/or race were missing or invalid, the Census Bureau: (1) assigned Hispanic origin and race from the 2010 Census response, from the ACS, from the Numident File, or from other federal sources; (2) assigned race from Hispanic origin, or vice versa; (3) allocated from another household member's response; or (4) allocated from the nearest neighbor hot deck.

Some respondents chose not to provide *relationship* information or provided invalid or inconsistent responses. Invalid could mean that multiple relationship categories were selected. Inconsistent would be when the relationship status did not match the age or sex of other members in the housing unit. When this happened, the Census Bureau: (1) assigned or allocated relationship to maintain consistency with other household members; (2) assigned relationship from the Census Bureau's Kidlink file—which contains SSA information and links parent-child relationships; or (3) allocated a value using the nearest neighbor hot deck imputation.

If housing *tenure status* (owner or renter) was missing or invalid (e.g., when multiple values were selected), the Census Bureau: (1) assigned tenure

[10] "Hot deck" imputation is a method whereby records are sorted geographically (state, county, census tract, block group, block), reported values from each record are input into a matrix of relevant variables for the variable requiring imputation, and the most recent reported value in the matrix is used to fill in the record with a missing value. The method is dynamic (although a donor may provide imputed values for more than one record) and takes advantage of correlations among values for people living in close proximity.

Copyright National Academy of Sciences. All rights reserved.

Table C.5 Differences in Numbers of Write-Ins for the Race and Hispanic-Origin Questions, 2010 and 2020 Censuses

Census/Question	2010 Write-Ins	2020 Write-Ins	Difference (2020 – 2010)
Race	37.7 million	335.5 million	297.8 million
Ethnicity (Hispanic origin)	17.1 million	15.0 million	−2.1 million

SOURCE: Jones et al. (2021:Slide 9).

using federal administrative records sources, such as information from the U.S. Department of Housing and Urban Development, or commercial tax and deed information; (2) allocated from the nearest neighbor hot deck using administrative records values as a nearness measure; or (3) allocated from the nearest neighbor hot deck.

Detailed vacancy status was typically collected by enumerators during the NFRU operation when a housing unit was identified as vacant, but housing units could also be identified as vacant using high-quality administrative records and through Count Imputation. If detailed vacancy (e.g., seasonally vacant) was missing, the Census Bureau allocated a value using nearest neighbor hot deck imputation.

Finally, if all characteristics information was missing for all persons in an occupied housing unit, such as in "population count only" cases, and when the housing unit status was count imputed, the Census Bureau allocated characteristics from administrative records associated with the housing unit when such records were available. Otherwise, the Census Bureau allocated using entire neighboring households from nearest neighbor hot deck imputation. These cases were termed whole-person imputations.

C.8.4 Changes in Format and Coding of the 2020 Race and Ethnicity Questions

Chapter 10 documents significant changes in the format of the race and ethnicity questions in 2020—most important, adding write-in spaces under the White and Black checkboxes—and in the coding of race and ethnicity responses. The greater opportunity for the provision of write-in responses resulted in very large increases in the number of write-in responses for race in the 2020 Census compared to the 2010 Census. However, this was not the case for write-ins for ethnicity—see Table C.5.

Copyright National Academy of Sciences. All rights reserved.

C.8.5 Disclosure Protection and Tabulation

The 2020 Census used a new DAS on publicly released data products, which has had major adverse impacts on the timeliness, quality, and utility of those products compared with earlier censuses. Many data products have just been released or are not yet released as of this writing. Chapter 11 describes the 2020 DAS and its implementation for the various 2020 data products.

C.9 2020 CENSUS EVALUATION

C.9.1 Coverage Evaluation

The Census Bureau traditionally uses two methods evaluating the completeness of the population count in total and for population groups and some geographies (e.g., states). These methods are Demographic Analysis (DA) and the Post-Enumeration Survey (PES). See Chapter 4 for details on each method and comparisons with the 2010 Census coverage measurement.

C.9.2 Census Program for Evaluations and Experiments

Like several of its predecessors, the 2020 Census Program for Evaluations and Experiments (CPEX) program combined experiments conducted during or shortly after the census, evaluations and assessments of the census, and quality profiles. *Census experiments* are generally intended to be forward looking, to identify techniques and methods that might prove beneficial for subsequent censuses. *Census evaluations* are analyses that interpret and synthesize the effectiveness and efficiencies of census components and their impact on data quality and coverage. Evaluations go beyond how successful (in terms of budget and deadline) the performance of a component process was. These studies can delve into the range of performances of a component process for various demographic groups, types of geographies, or other subsets of the population (e.g., owner/renter). *Census assessments* document final volumes, rates, and costs for individual operations or processes using data from production files and information collected from debriefings. They are more focused and more descriptive, as opposed to analytical, than are evaluations. *Quality profiles* present the results of quality-assurance programs for census operations.

Below, we catalog the titles of experiments, evaluations, and quality profiles from the 2010 and 2020 CPEX, providing release dates and scheduled release dates, as applicable. Publication of reports from the 2020 CPEX is running behind that for 2010. The titles and timelines are derived from the evaluations landing webpage for the two censuses.[11]

[11]See https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/evaluate/eae.html for 2020 and https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/evaluate/eae.2010.html for 2010.

Copyright National Academy of Sciences. All rights reserved.

- **Experiments**
  - *2020—3 reports*
    - Real-Time 2020 Administrative Records Simulation—5/5/2023
    - Optimization of Self-Response in the 2020 Census—*By December 2023*
    - Extending the Decennial Census Environment to the Mailing Materials—4/10/2023
  - *2010—6 reports*
    - Race and Hispanic Origin Alternative Questionnaire Experiment—2/28/2013
    - 2010 Census Nonresponse Followup Contact Strategy Experiment Report—1/30/2012
    - 2010 Census Confidentiality Notification Experiment Report—1/12/2012
    - 2010 Census Deadline Messaging Compressed Mailing Schedule Experiment—12/14/2011
    - The Paid Advertising Heavy-Up Experiment—4/24/2011
    - 2010 Census AQE: Census 2000 Form Replication Panel—1/10/2011
- **Evaluations**
  - **2020—12 reports**
    - Analysis of Census Internet Self-Response Paradata by Language—*By September 2024*
    - Research on Historically Undercounted Populations: Non-English Speakers and Complex Household Residents including Undercount of Children—*By September 2024*
    - Evaluation of Reengineered Address Canvassing Operation—*By September 2024*
    - Administrative Records Dual System Estimation—*By March 2024*
    - 2020 Census Communications Campaign Evaluation: Census Mindset Measures Before and After the Campaign—*By September 2024*
    - Privacy and Confidentiality Concerns—*By September 2024*
    - Matching 2018 Census Barriers, Attitudes, and Behaviors Study (CBAMS) Survey Sample to the 2020 Census—*By December 2023*
    - 2020 Census Tracking Survey—*By December 2023*
    - Group Quarters Advance Contact: Refining Classification of College or University Student Housing—*By December 2023*

Copyright National Academy of Sciences. All rights reserved.

    - Investigating Digital Advertising and Online Self-Response—*By December 2023*
    - Comparing 2019 Census Test and 2020 Census Self-Response Rates to Estimate the "Decennial Environment"—1/23/2023
    - 2020 Census Quantitative Creative Testing Final Report—11/16/2022

  - **2010—12 reports**
    - Comparative Ethnographic Studies of Enumeration Methods and Coverage—3/29/2013
    - Evaluation of Address List Maintenance Using Supplemental Data Source—2/20/2013
    - Administrative Records Use for Coverage Problems Evaluation Report—2/13/2013
    - Evaluation of Address Frame Accuracy and Quality—1/14/2013
    - Observing Census Enumeration of Non-English Speaking Households—12/6/2012
    - Behavior Coding Report Coverage Measurement Person Interviews—12/4/2012
    - 2010 Census Match Study—11/16/2012
    - Alternative Coverage Followup Questions and Design Evaluation—10/16/2012
    - Effectiveness of Unduplication Evaluation—9/25/2012
    - Investigation of the Methods to Evaluate the Coverage of Group Quarters—9/25/2012
    - 2010 Census Evaluation to Assess CCM Search Area & Address List Rules—9/20/2012
    - Global Positioning System Evaluation—9/12/2012
    - 2010 Census Avoid Followup Evaluation—9/7/2012

- **Operational Assessments**
  - **2020—46 reports**
    - 2 Assessment Reports (Scheduled)—*By June 2025*
    - 2 Assessment Reports (Scheduled)—*By December 2024*
    - 2 Assessment Report (Scheduled)—*By September 2024*
    - 12 Assessment Reports (Scheduled)—*By March 2024*
    - 17 Assessment Reports (Scheduled)—*By December 2023*
    - 2 Assessment Reports (Scheduled)—*By September 2023*
    - 2 Assessment Reports: Count Review and Post-Enumeration Survey Initial Housing Unit Matching Assessment—August 8 and 17, 2023

Copyright National Academy of Sciences. All rights reserved.

    - 2 Assessment Reports: Item Nonresponse and Imputation; Census Questionnaire Assistance (CQA)—February 7 and 9, 2023
    - 1 Assessment Reports: Post-Enumeration Survey Independent Listing—January 24, 2023
    - 2 Assessment Reports: Local Update of Census Addresses (LUCA); Archiving—November 9 and 28, 2022
    - 2 Assessment Reports: In-Field Address Canvassing; In-Office Address Canvassing—August 18, 2022
  - **2010—59 reports**
    - 1 Assessment Report—January 2013
    - 52 Assessment Reports—January–December 2012
    - 6 Assessment Reports—March–December 2011
- **Quality Profiles/Quality Assurance Reports**
  - **2020—3 reports**
    - Nonresponse Followup Quality Assurance Results—8/7/2023
    - Update Leave Quality Assurance Results—12/7/2022
    - In-Field Address Quality Assurance Results—9/14/2022
  - **2010—6 reports**
    - 2010 Census Enumeration at Transitory Locations Reinterview Operation—6/26/2012
    - 2010 Census Address Canvassing Quality Profile—6/21/2012
    - 2010 Census: Nonresponse Followup Quality Profile—3/20/2012
    - 2010 Census: Update/Leave Quality Profile—12/12/2011
    - 2010 Census Coverage Measurement Independent Listing Quality Profile—11/30/2011
    - 2010 Census Coverage Measurement Initial HU Followup Quality Profile—11/16/2011

Copyright National Academy of Sciences. All rights reserved.

– D –

Extensions of Census Coverage Evaluations

D.1 ESTIMATES OF CENSUS ERROR FROM THE POST-ENUMERATION SURVEY

The primary source of detailed information beyond Demographic Analysis on census enumeration errors is the Post-Enumeration Survey (PES). Briefly (see Chapter 4 for more detail and references), the PES is an independent survey of households in a sample of blocks (the P-sample), which is matched to census enumerations (referred to as the E-sample) in the same blocks. The U.S. Census Bureau, using capture-recapture or dual-system estimation, estimates net undercoverage in the census for various domains (e.g., age, sex, race, ethnicity, housing tenure). The Census Bureau also uses the PES to estimate the contributions of various operations to errors—finding, for example, that proxy enumerations generate more erroneous enumerations and whole-person imputations than other modes. A further analysis of census errors by component (e.g., duplications, omissions) would partition the population in the PES sample into two groups—those addresses, households, or individuals for which attempts at enumeration resulted in a census component enumeration error and those that did not. This analysis, in turn, could help assess the quality of each census process in enumerating the population and provide a basis for considering specific processes for improvement in the future. In this appendix, we outline such an approach.

In making use of the PES in this way, one would consider a large number of households or individuals with the binary attribute (proper or improper

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

enumeration).[1] Then, one would utilize various indicator variables identifying the type of living situation, the demography and geography of the household or individual, and the census processes used for that household or individual as predictors, to try to discriminate or predict proper or improper enumerations. The hope is that some variables shown to have predictive power will identify procedures that could be modified for 2030.

Considerable care is needed in interpreting these associations because selection effects challenge causal validity. For a basic example, consider the subset of data collected in an Internet First domain versus an Internet Choice domain, in which households in both instances elected to respond via paper questionnaire. In both cases the collection mode is "form", but the quality may well vary between the two because one rejected the internet, while the other chose the form. How to extract valid action-worthy conclusions from this basic example and from more complicated examples needs to be addressed.

Keeping the foregoing in mind, estimating associations is very worthwhile. Statistical models relevant in this regard include discriminant analysis and classification trees. Analyses can be broadened from the individual to, for example, geographic aggregates with the relative frequency of enumeration error as the dependent variable. For such aggregate models, groups of census tracts offer a great advantage over other aggregates such as states and counties, because tracts are more homogeneous with respect to census processes and living situations. Furthermore, using such a unit of aggregation allows one to draw on information from the American Community Survey (ACS), which can offer additional predictors, such as the degree of access to broadband internet connections, which could help identify reasons for the frequency of Internet Self-Response.

National Research Council (2009:Ch. 5) provides useful suggestions for the proposed analysis, with a focus on answering this general question: "Which census processes are associated with a substantially increased rate of erroneous enumerations, duplications, omissions, or enumerations in the wrong location?" Specific questions follow, including types of housing units that were missed or duplicated more often than others and types of people who were missed, erroneously enumerated, duplicated, or counted in the wrong locations more often than others. Statistical modeling could help answer such questions, with the expectation that the findings would identify pathways for census improvement in 2030.

The modeling approach outlined here differs somewhat from that described by the National Research Council (2009). That report suggested conducting the analysis at the level of census domains. The proposed approach also includes analysis at the level of individuals or households.

[1]The outcome could be expanded to a more refined metric.

Copyright National Academy of Sciences. All rights reserved.

D.1.1 Candidate Attributes

Candidates for *dependent variables* include indicators of four main types of error:[2]

- Whether a P-sample enumeration is a census omission—refinements would include indicator variables as to whether the omission is: (a) a within-household omission, for which others in the same housing unit were enumerated; (b) a whole-household omission in a building included in the Master Address File (MAF); or (c) an entirely missed housing unit.
- Whether an E-sample enumeration is a census duplicate—refinements would include indicator variables that denote whether the situation is: (a) a whole-household duplicate; or (b) a duplication of an individual in a household in which others were counted only once. It could also be useful to distinguish between indicators of duplications of individuals in non-group quarters residences with people in various types of group quarters (GQ) residences (college/university student housing, nursing homes, correctional facilities, military barracks), or indicators of duplications of individuals in standard residences with someone in a seasonal residence, or indicators of duplications of nonmovers with a recent mover.
- Whether an E-sample enumeration is erroneous—refinements would include whether the erroneous enumeration is: (a) a fictitious person; (b) someone who was born after Census Day; (c) someone who died before Census Day; or (d) a visitor.
- Whether an E-sample enumeration is in the wrong location—refinements would include whether the wrong location is due to: (a) a geocoding error; (b) a recent move; or (c) a person counted at a Census Day residence that is not their usual residence.

Other dependent variables could include missing item responses, item responses in error, population-count-only households, and others.

Candidates for *predictor variables* are identified by the National Research Council (2009:123):

> To understand which subset of individuals and housing units are more frequently subject to coverage errors of the four indicated types, and to understand what census processes contributed to those errors, it is necessary to focus on *predictors that distinguish between individuals and housing units that are likely to have different interactions with census enumeration processes, as well as predictors that indicate the census processes*

[2]Some of these analyses might be improved through use of a non-dichotomous loss structure reflecting errors of differing gravity. The most natural example is counting people in the wrong location, which has the natural ordering of counts associated with locations that only become correct at some higher level of aggregation being more serious errors than those for locations that become correct at some lower level of aggregation.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

> *that were used to attempt to enumerate those individuals and housing units* [emphasis added].
>
> We caution that the set of census component processes that one might want to represent is much richer than the discussion here would suggest. To get some sense of the complexity, there will be dozens of different questionnaires used in the 2010 Census (to account for different forms of delivery, foreign languages, and other factors). Given the size of the postenumeration surveys (PES) used to date, the sample sizes of the coverage measurement program are unlikely to support analysis of the rarer component processes used in conjunction with the more detailed subsets of the population. Therefore, some compromise between full representation and parsimony in modeling must be struck.
>
> We recognize that some of the covariates suggested for use in statistical modeling of the frequency of various components of coverage error may not be routinely available given the planned data collection in the P-sample blocks. For instance, we discuss below the possibility of determining whether someone in a household telephoned for questionnaire assistance and whether that was associated with a higher or lower rate of census coverage error of some type. Such information has not been collected in previous coverage measurement programs and would therefore not be available to modelers. . . . If this information will not be available, we hope that this discussion motivates a revision of plans to make more predictors available to support the models.

The types of predictors that would be useful for 2030 are indicators of: (1) the area of interest; (2) census processes; (3) the degree of enumeration error; and (4) context.

Indicators of the area of interest could include:

- Type of enumeration area and other features that identify the local geography and the types of housing units in the area—candidate variables include the quality of the mailing list and whether the housing units have unique identifiers; the frequency of small multiunit residences, the rates of new construction and recent demolition; and the degree of mixing of residential and business establishments.
- Housing unit variables—candidate variables include whether the household itself is newly constructed or is part of a small multiunit building and whether the housing unit is part of a GQ and, if so, what type.
- People's demographic characteristics—candidate variables include indicators for demographic groups that are historically subject to varying degrees of net census undercoverage (e.g., ages 18–22 are associated with increased chances of duplication and ages 0–4 with increased chances of omission).
- Relationships of residents—candidate variables include whether the household includes one or more unrelated people or more than one nuclear family.

Copyright National Academy of Sciences. All rights reserved.

Indicators of census processes could include:

- Results of the MAF building process—four types of coverage error can stem from mistakes in MAF building: (a) if a nonresidential unit is included in the MAF; (b) if a housing unit is included twice (at different street addresses); (c) if a housing unit is omitted (which can happen in multiple ways since the building a residence is found in could be included but the residence omitted, or the entire building could be missed); and (d) if an address is geocoded in the wrong location. While none of these errors necessarily results in a census coverage error (since field work can remedy such mistakes) such errors are likely associated with increased frequency of coverage error. Therefore, indicator variables for these types of error are obvious variables for inclusion in such models. Further, it might be useful to use such indicator variables themselves as dependent variables in models that attempt to find out why such errors occurred.
- Variables associated with self-response—indicator variables could include that a foreign language questionnaire was requested or some other contact was made with telephone questionnaire assistance, some degree of item nonresponse occurred on the returned questionnaire, some of the response needed to be keyed, or the response did not use the provided census identification number (Non-ID response).
- Nonresponse Followup (NRFU)-associated variables—indicator variables could include the number of nonresponse attempts needed and whether the ultimate enumeration was through a proxy respondent or through whole-household imputation.
- Variables associated with other modes of enumeration—indicator variables could include seasonal residence in an Update Leave or Update Enumerate area.

Indicators of degree of enumeration error could include:

- Variables associated with enumerator training and turnover rates.
- Variables from the quality-control checks of enumerators' work.

Indicators of contextual factors could include:

- Variables associated with neighborhood characteristics, such as the percentage of people in an area that own their own residences, the local mail return rate, the local crime rate, and the health of the local economy.

D.1.2 Implementation Issues

As noted in National Research Council (2009:126):

> It would be premature at this point to suggest the precise form of the statistical models that could be used for this application, but since most of the dependent variables are dichotomous, logistic regression, discriminant

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

analysis, and classification trees (including random forests) are obvious models to consider. Moreover, it is very likely that given the complexity of the underlying phenomena being modeled, focusing on models with predictors restricted to the poststratification variables from 2000 will be unsatisfactory. Instead, what is needed is a representation of the complexity of the situation, involving characteristics of households, housing units, census processes, enumerator's performance, and interactions among these variables. In addition, the Census Bureau should also examine the possibility of using separate regression models for separate geographic domains.

D.1.3 PES Redesign

While the usual structure of the PES could be continued in 2030 for both the traditional net and gross coverage error estimation and the proposed discriminant analysis, other designs could provide greater utility in this regard. One way to make the PES sample useful for statistical modeling in addition to the traditional estimation would be to oversample areas that were likely to experience problems with census enumeration. For instance, three-quarters of the usual PES design could be structured as it was in the past, with one-quarter of the sample devoted to oversampling areas in which problems were anticipated. The hope would be to get more examples of improper enumerations for the logistic regression or classification tree in addition to the many instances of no enumeration problems. For example, additional sample could be added in immigrant communities, which have a higher rate of omissions; seasonal housing communities, which have a higher rate of duplications; and the like. This design would not add a substantial amount to the estimated error of the coverage estimates for large domains. However, given that the causal structure of census enumeration errors is not fully understood, deciding what types of areas to oversample is not completely obvious.

D.2 RESEARCH TO GENERATE TESTABLE IDEAS TO REDUCE THE DIFFERENTIAL NET UNDERCOUNT IN 2030

D.2.1 The Problem

Differential net undercount is a continuing problem in the U.S. decennial census. Groups that are historically *undercounted* include Black people, Hispanic people, American Indian or Alaska Native people, renters, young children, young men from various demographic groups, areas with low income and low levels of education, areas with large numbers of recent immigrants, and remote rural areas. Other groups are historically *overcounted*, including older people, homeowners, and White people (also Asian people in 2020). There are many reasons for undercount, and the Census Bureau has successfully reduced

Copyright National Academy of Sciences. All rights reserved.

net undercount rates for the total population and population groups since 1950. However, the differential net undercount (e.g., between Black people and all others) remains stubbornly high at 3–4 percentage points. Moreover, the 2020 Census experienced higher net undercount rates for disadvantaged groups than were seen in 2010, and a wider differential. These differential undercounts are consequential for critical uses of census data that involve allocating resources from a fixed pie, including redistricting and federal fund allocation, and they affect other important uses as well.

Identifying strategies that could reduce differential net undercount has suffered from a lack of granular data for analysis. Demographic Analysis does not provide subnational estimates or estimates for race groups other than Black and All Other Races. The PES provides race and ethnicity detail at the national level and total state estimates, but not substate estimates. Administrative records vary in their completeness of population coverage as well as race and ethnicity detail (their use for race and ethnicity imputation, as well as other characteristics, such as age, should be studied).

D.2.2 Proposal for a Data Linkage Study Focused on Undercounts

The proposed study would link ACS individual records for 2019, 2020, and 2021 to the 2020 Census and to as many kinds of administrative records as possible and not only those used nationwide in the 2020 Census. The goal would be to tease out patterns and characteristics that suggest strategies that could improve the census count for some groups in some areas.[3] The study would need to be carried out by Census Bureau staff or, perhaps better, by a consortium of researchers working with census staff and people with local and group knowledge at a Federal Statistical Research Data Center. It would be ideal if the study could be completed within 18–24 months and presented to the Census Bureau and stakeholder leadership with recommendations for actionable ideas to pursue (such as obtaining authorization for a particular type of administrative record or illustrating the extent of positive effects on self-response of universal affordable broadband).

The ACS has response and coverage problems itself; however, the hypothesis is that, just as in the PES, matching the ACS with the census and administrative records would find some addresses missed in the ACS that are found in the census and vice versa. Also, some households would match but be missing one or more household members in the ACS or the census. Similarly, various administrative records would include or omit people found in the ACS

[3]To the argument that there should not be separate methods for specific areas or groups in the census, a response is that the census has always had a range of methods—for example, Update Enumerate in remote areas. Moreover, the goal is equity in the count, which given long-lasting differential net undercounts, necessarily means additional and/or different methods directed in a targeted way to try to reduce the differential.

Copyright National Academy of Sciences. All rights reserved.

Table D.1 Hypothetical Example of How a Specific 5-Person Household Would Appear in Various Data Sources

	Adult		Citizen Child		
Data Source	Legal Resident	Undocumented	Age 12	Age 9	Age 3
2020 Census	Yes	No	Yes	Yes	No
2019 ACS	Yes	No	Yes	No	No
Birth Records	No	No	Yes	Yes	Yes
Earnings Records (SSA, LEHD)	Yes	Yes	No	No	No
SNAP Records	Yes	No	Yes	Yes	Yes
SLDS (School) Records	No	No	Yes	Yes	Yes

NOTES: ACS, American Community Survey; SSA, Social Security Administration; LEHD, Longitudinal Employer-Household Dynamics; SLDS, State Longitudinal Data System; SNAP, Supplemental Nutrition Assistance Program.

SOURCE: Panel generated.

and/or census and vice versa. The hope would be to obtain more complete representation of households with information about which sources and combinations of sources are better for including households in the census. The extensive demographic detail available in the ACS for individual respondents and for neighborhoods would enable exploration of a number of scenarios and possibilities.

As an example, consider a household of a legal resident adult, an undocumented adult, and three citizen children. The household might show up as depicted in Table D.1 in various sources, suggesting ways to count people who would otherwise be missing in the census.

D.2.3 Feasibility

Is such a study feasible in terms of numbers of available cases? It appears so. Each year, the ACS has 3.5 million addresses in sample that typically yields about 2 million completed interviews. For 2019, 2020, and 2021, the ACS had 9.9 million addresses in sample and 5.4 million completed interviews. The average tract has about 1,650 addresses, of which about 120 addresses would be in sample. For analysis purposes, a possible plan could be to identify the lowest one-quarter of census tracts in Self-Response—about 21,000 tracts—in states for which the Census Bureau has Supplemental Nutrition Assistance Program records (setting aside tracts in remote areas, on American Indian or Alaska Native reservations, and the like for separate attention). To generate sufficient cases for analysis, a plan could be to combine the 21,000 tracts into groups of, say, five, which would generate 4,200 analysis units with 600 addresses in each

Copyright National Academy of Sciences. All rights reserved.

unit. The groupings would include tracts with similar characteristics, such as race/ethnicity, broadband availability, etc. Records would be matched (in a protected setting) on an individual basis. There would be error in matching, as there is in the PES, but matching at a microlevel on the scale proposed (about 2.5 million addresses) and with as many record sources as possible could hopefully identify a number of avenues for testing in the lead-up to the 2030 Census. While there is no guarantee of success, without such a study it seems unlikely that new evidence-based ideas would emerge that would prove fruitful in testing.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

– E –

2020 Census Group Quarters Definitions and Type Code List

This appendix reprints, in full, the Census Bureau's listing of group quarters types posted at https://www2.census.gov/programs-surveys/decennial/2020/program-management/pcgqr/2020-pcgqr-participants-guide-english.pdf#page=34 made available to participants in the 2020 Census Post Census Group Quarters Review program. The only difference is formatting in paragraphs rather than an extended table.

GROUP QUARTERS

A group quarters is a place where people live or stay, in a group living arrangement that is owned or managed by an entity or organization providing housing and/or services for the residents. These services may include custodial or medical care as well as other types of assistance, and residency is commonly restricted to those receiving these services. This is not a typical household-type living arrangement. People living in group quarters are usually not related to each other.

Group quarters include such places as college residence halls, residential treatment centers, skilled-nursing facilities, group homes, military barracks, correctional facilities, and workers' dormitories.

Copyright National Academy of Sciences. All rights reserved.

INSTITUTIONAL GROUP QUARTERS

Facilities that house those who are primarily ineligible, unable, or unlikely to participate in the labor force while residents.

Correctional Facilities for Adults

- **Federal Detention Centers** (101): Stand alone, generally multi-level, federally operated correctional facilities that provide "short-term" confinement or custody of adults pending adjudication or sentencing. These facilities may hold pretrial detainees, holdovers, sentenced offenders, and Immigration and Customs Enforcement (ICE) inmates, formerly called Immigration and Naturalization Service (INS) inmates. These facilities include: Metropolitan Correctional Centers (MCCs), Metropolitan Detention Centers (MDCs), Federal Detention Centers (FDCs), Bureau of Indian Affairs Detention Centers, ICE Service Processing Centers, and ICE contract detention facilities.
- **Federal and State Prisons** (102 Federal Prisons; 103 State Prisons): Adult correctional facilities where people convicted of crimes serve their sentences. Common names include: prison, penitentiary, correctional institution, federal or state correctional facility, and conservation camp. The prisons are classified by two types of control: (1) "federal" (operated by or for the Bureau of Prisons of the Department of Justice) and (2) "state." Residents who are forensic patients or criminally insane are classified on the basis of where they resided at the time of enumeration. Patients in hospitals (units, wings, or floors) operated by or for federal or state correctional authorities are counted in the prison population. Other forensic patients will be enumerated in psychiatric hospital units and floors for long-term non-acute patients. This category may include privately operated correctional facilities.
- **Local Jails and Other Municipal Confinement Facilities** (104): Correctional facilities operated by or for counties, cities, and American Indian and Alaska Native tribal governments. These facilities hold adults detained pending adjudication and/or people committed after adjudication. This category also includes work farms and camps used to hold people awaiting trial or serving time on relatively short sentences. Residents who are forensic patients or criminally insane are classified on the basis of where they resided at the time of enumeration. Patients in hospitals (units, wings, or floors) operated by or for local correctional authorities are counted in the jail population. Other forensic patients will be enumerated in psychiatric hospital units and floors for long-term non-acute patients. This category may include privately operated correctional facilities.

Copyright National Academy of Sciences. All rights reserved.

- **Correctional Residential Facilities** (105): These are community-based facilities operated for correctional purposes. The facility residents may be allowed extensive contact with the community, such as for employment or attending school, but are obligated to occupy the premises at night. Examples are halfway houses, restitution centers, and prerelease, work release, and study centers.
- **Military Disciplinary Barracks and Jails** (106): Correctional facilities managed by the military to hold those awaiting trial or convicted of crimes.

Juvenile Facilities

- **Group Homes for Juveniles (non-correctional)** (201): Includes community-based group living arrangements for youth in residential settings that are able to accommodate three or more clients of a service provider. The group home provides room and board and services, including behavioral, psychological, or social programs. Generally, clients are not related to the care giver or to each other. Examples are maternity homes for unwed mothers, orphanages, and homes for abused and neglected children in need of services. Group homes for juveniles do not include residential treatment centers for juveniles or group homes operated by or for correctional authorities.
- **Residential Treatment Centers for Juveniles (non-correctional)** (202): Includes facilities that primarily serve youth that provide services on-site in a highly structured live-in environment for the treatment of drug/alcohol abuse, mental illness, and emotional/behavioral disorders. These facilities are staffed 24-hours a day. The focus of a residential treatment center is on the treatment program. Residential treatment centers for juveniles do not include facilities operated by or for correctional authorities.
- **Correctional Facilities Intended for Juveniles** (203): Includes specialized facilities that provide strict confinement for its residents and detain juveniles awaiting adjudication, commitment or placement, and/or those being held for diagnosis or classification. Also included are correctional facilities where residents are permitted contact with the community, for purposes such as attending school or holding a job. Examples are residential training schools and farms, reception and diagnostic centers, group homes operated by or for correctional authorities, detention centers, and boot camps for juvenile delinquents.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Nursing Facilities/Skilled-Nursing Facilities

Nursing Facilities/Skilled-Nursing Facilities (301): Includes facilities licensed to provide medical care with seven day, twenty-four hour coverage for people requiring long-term non-acute care. People in these facilities require nursing care, regardless of age. Either of these types of facilities may be referred to as nursing homes.

Other Institutional Facilities

- **Mental (Psychiatric) Hospitals and Psychiatric Units in Other Hospitals** (401): Includes psychiatric hospitals, units and floors for long-term non-acute care patients. The primary function of the hospital, unit, or floor is to provide diagnostic and treatment services for long-term non-acute patients who have psychiatric-related illness. All patients are enumerated in this category.
- **Hospitals with Patients Who Have No Usual Home Elsewhere** (402): Includes hospitals if they have any patients who have no exit or disposition plan, or who are known as "boarder patients" or "boarder babies." All hospitals are eligible for inclusion in this category except psychiatric hospitals, units, wings or floors operated by federal, state or local correctional authorities. Patients in hospitals operated by these correctional authorities will be counted in the prison or jail population. Psychiatric units and hospice units in hospitals are also excluded. Only patients with no usual home elsewhere are enumerated in this category.
- **In-Patient Hospice Facilities** (403): Includes in-patient hospice facilities (both free-standing and units in hospitals) that provide palliative, comfort, and supportive care for the terminally ill patient and their families. Only patients with no usual home elsewhere are tabulated in this category.
- **Military Treatment Facilities with Assigned Patients** (404): These facilities include military hospitals and medical centers with active duty patients assigned to the facility. Only these patients are enumerated in this category.
- **Residential Schools for People with Disabilities** (405): Includes schools that provide the teaching of skills for daily living, education programs, and care for students with disabilities in a live-in environment. Examples are residential schools for the physically or developmentally disabled.

NONINSTITUTIONAL GROUP QUARTERS

Facilities that house those who are primarily eligible, able, or likely to participate in the labor force while residents.

Copyright National Academy of Sciences. All rights reserved.

College/University Student Housing

College/University Student Housing (501): Includes residence halls, and other buildings, including apartment-style student housing, designed primarily to house college and university students in a group living arrangement either on or off campus. These facilities are owned, leased, or managed either by a college, university, or seminary, or by a private entity or organization. Fraternity and sorority housing recognized by the college or university are included as college student housing. Students attending the U.S. Naval Academy, the U.S. Military Academy (West Point), the U.S. Coast Guard Academy, and the U.S. Air Force Academy are counted in military group quarters.

Military Quarters

Military Quarters (601 Military Quarters; 602 Military Ships): These facilities include military personnel living in barracks (including "open" barrack transient quarters) and dormitories and military ships. Patients assigned to Military Treatment Facilities and people being held in military disciplinary barracks and jails are not enumerated in this category. Patients in Military Treatment Facilities with no usual home elsewhere are not enumerated in this category.

Other Noninstitutional Facilities

- **Emergency and Transitional Shelters (with Sleeping Facilities) for People Experiencing Homelessness** (701): Facilities where people experiencing homelessness stay overnight. These include:
  1) shelters that operate on a first-come, first-serve basis where people must leave in the morning and have no guaranteed bed for the next night;
  2) shelters where people know that they have a bed for a specified period of time (even if they leave the building every day); and
  3) shelters that provide temporary shelter during extremely cold weather (such as churches). This category does not include shelters that operate only in the event of a natural disaster.

  Examples are emergency and transitional shelters; missions; hotels and motels used to shelter people experiencing homelessness; shelters for children who are runaways, neglected or experiencing homelessness; and similar places known to have people experiencing homelessness.
- **Soup Kitchens, Regularly Scheduled Mobile Food Vans, and Targeted Non-Sheltered Outdoor Locations** (702 Soup Kitchens; 704 Regularly Scheduled Mobile Food Vans; 706 Targeted Non-Sheltered Outdoor Locations): Includes soup kitchens that offer meals organized as food service lines or bag or box lunches for people experiencing homelessness;

Copyright National Academy of Sciences. All rights reserved.

street locations where mobile food vans regularly stop to provide food to people experiencing homelessness; and targeted non-sheltered outdoor locations where people experiencing homelessness live without paying to stay. This would also include persons staying in pre-identified car, RV and tent encampments. Targeted non-sheltered outdoor locations must have a specific location description; for example, "the Brooklyn Bridge at the corner of Bristol Drive," "the 700 block of Taylor Street behind the old warehouse," or the address of the parking lot being utilized.

- **Group Homes Intended for Adults** (801): Group homes are community-based group living arrangements in residential settings that are able to accommodate three or more clients of a service provider. The group home provides room and board and services, including behavioral, psychological, or social programs. Generally, clients are not related to the care giver or to each other. Group homes do not include residential treatment centers or facilities operated by or for correctional authorities.
- **Residential Treatment Centers for Adults** (802): Residential facilities that provide treatment on-site in a highly structured live-in environment for the treatment of drug/alcohol abuse, mental illness, and emotional/behavioral disorders. They are staffed 24-hours a day. The focus of a residential treatment center is on the treatment program. Residential treatment centers do not include facilities operated by or for correctional authorities.
- **Maritime/Merchant Vessels** (900): Includes U.S. owned and operated flag vessels used for commercial or noncombatant government related purposes at U.S. ports, upon the sea, or on the Great Lakes.
- **Workers' Group Living Quarters and Job Corps Centers** (901): Includes facilities such as dormitories, bunkhouses, and similar types of group living arrangements for agricultural and non-agricultural workers. This category also includes facilities that provide a full-time, year-round residential program offering a vocational training and employment program that helps young people 16-to-24 years old learn a trade, earn a high school diploma or GED and get help finding a job. Examples are group living quarters at migratory farmworker camps, construction workers' camps, Job Corps centers and vocational training facilities.
- **Living Quarters for Victims of Natural Disasters** (903): These are temporary group living arrangements established as a result of natural disasters.
- **Other Noninstitutional Group Quarters** (904): Other noninstitutional group quarters

Copyright National Academy of Sciences. All rights reserved.

– F –

Additional Detail and Reference on Data Products and Disclosure Avoidance

F.1 EXPLANATION OF DIFFERENTIAL PRIVACY

Two formulations of differential privacy (DP), pure/approximate DP (the original approach) and zero-concentrated DP (zCDP), are outlined with sufficient mathematical formalism to communicate the concepts and identify control parameters. For additional details and application to the 2020 Census, see Abowd and Hawes (2023), Gong et al. (2022), and Special Issue 2 of the *Harvard Data Science Review* (2022). See also Shlomo (2023).

F.1.1 Pure/Approximate Differential Privacy

Let x and x' be two databases that differ by one record (e.g., x is the full data set and x' has one record deleted). For $\epsilon > 0$, $\delta \in [0, 1]$, $M(\cdot)$ a noise-infusing function that operates on a database, and E a summary of the output, the system is (ϵ, δ) differentially private if,

$$P\left(M(x) \in E\right) \leq \exp(\epsilon) \times P\left((M(x') \in E\right) + \delta \tag{F.1}$$

for all possible E. To see how this works, consider $\delta = 0$ (pure DP), then $\exp(\epsilon)$ controls the likelihood ratio of the outcome (E) produced from x relative to x'. That is, Equation F.1 controls the multiplicative change in the probability of a reported measurement based on a full dataset and a dataset

Copyright National Academy of Sciences. All rights reserved.

with one respondent removed, and the probability of identification/disclosure is larger for a larger likelihood ratio. The protection guarantee is probabilistic and relative/comparative, not absolute. The parameter ϵ is the "privacy loss budget." Setting $\epsilon = 0$ constrains the likelihood ratio to equal 1.0 (no transmitted information, but complete protection). Increasing ϵ without bound allows communication of full information with no protection.

The approach provides *global protection* for all possible outcomes E without regard to the consequences of disclosure or the particular information available to potential intruders. Setting $\delta > 0$ (approximate DP) introduces the possibility of violating the pure DP likelihood ratio bound, $\exp(\epsilon)$, to accommodate low-prevalence Es. The U.S. Census Bureau adopted $\epsilon = 19.61$ for the 2020 Redistricting File, 17.14 for the person tables and 2.47 for the housing unit tables.[1] The amount of disclosure protection actually realized with such high parameter values is harder to know.

If k multiple, independent queries are posed to the posted data, the overall privacy loss budget equals $\epsilon_1 + \cdots + \epsilon_k$, the sum of the individual privacy loss budgets. So, maintaining an overall privacy loss budget of ϵ^* requires individual budgets that average $\epsilon^{*/k}$, which can be very small, allowing very little information to "leak through."

The approach is similar in spirit to a minimax statistical procedure, one that bounds the maximum risk for all states of nature. Minimax "pays" for this protection by reducing the benefit, sometimes considerably, for what might be the most important and likely underlying states of nature.

F.1.2 Zero-Concentrated Differential Privacy

For details, see Abowd et al. (2022). For $\rho \geq 0$ and $\alpha > 1$, a randomized mechanism satisfies ρ-zero-concentrated differential privacy ((α, ρ)-zCDP) if

$$D_\alpha \{M(x) \parallel M(x')\} \leq \rho\alpha, \text{ for all } \alpha > 1, \tag{F.2}$$

where D_α is the Renyi divergence of order α (a measure of distance between two probability distributions). Here, rather than controlling disclosure for all output subsets (E), the criterion is to ensure that the full distributions of output data for the two datasets are "close." As for the pure/approximate approach in Equation F.1, control is global with no particular attention to important aspects of the difference in distributions, and the privacy loss budget, controlled by (α, ρ), increases with the number of queries.[2]

[1]This text was revised after release of the prepublication version of the report to cite Census Bureau press release language (https://www.census.gov/newsroom/press-releases/2021/2020-census-key-parameters.html) and to strike a previous assertion to implications for the likelihood ratio associated with confidential information disclosure.

[2]Enhancements reduce $k \cdot \epsilon$ to $\sqrt{k \cdot \epsilon}$.

Copyright National Academy of Sciences. All rights reserved.

F.1.3 Census Bureau Adaptations of Differential Privacy to the 2020 Census Context

Differential privacy is at the core of the Census Bureau's 2020 Disclosure Avoidance System (DAS).[3] However, the Census Bureau needed to adapt the mathematically posed DP approaches to the realities of the 2020 Census. Realities include a wide variety of geographic and demographic domains that require different protection versus information trade-offs, a very large and to a degree unknowable set of future queries on the posted data, and face-validity constraints. The Census Bureau's approach for the 2020 Redistricting and Demographic and Housing Characteristics (DHC) Files was to create a "spine" to which most (not all) of these domains attached and use a TopDown Algorithm (TDA) to infuse noise with varying disclosure budgets for different domains (on-spine, off-spine, American Indian or Alaska Native spine).[4] Since the privacy loss budget increases with the number of queries, domain-specific budgets, indexed by (ϵ, δ) or (α, ρ), need to be very small to bound overall disclosure risk. However, a small budget produces low information in posted data, especially for small-population domains.

To the Census Bureau's credit, it held workshops (see National Academies of Sciences, Engineering, and Medicine, 2020, 2023) and in other ways solicited user feedback and from these inputs increased the privacy loss budget for some domains. The increases improved the usefulness of posted data with a consequent increase in identification risk. It remains to be seen whether these increased budgets have unacceptably increased exposure risk.

In one modification of generic DP, the Census Bureau defined several "invariants," for example, that a state total in posted data match the underlying, collected total. Invariants are important for face validity and some legal requirements, but they increase disclosure risk. In a similar vein, posted counts need to be non-negative, and noise infusion for a small count can produce a negative value. So, a non-negative constraint, implemented by multipass least squares to minimize discrepancies, was applied to all counts. Negative counts were transformed to non-negative values and to honor invariants, noise-infused positive counts were reduced. These adjustments induce positive bias for small population domains and some negative bias for larger population domains.

While the mathematics underlying DP are transparent to the mathematically inclined, modifications such as the spine, invariants, and the non-negativity constraint make the full process translucent at best, possibly bordering on

[3]See Abowd (2021b); Abowd et al. (2022); Abowd and Hawes (2023); Gong et al. (2022); U.S. Census Bureau (2021e); and *Harvard Data Science Review: Special Issue on Data Privacy and the 2020 Census*.

[4]Different algorithms are being developed for the DDHC and S-DHC files, but their properties are not yet well documented. A March 14, 2022, white paper by Tumult Labs describes the SafeTab-P algorithm for DDHC-A (available at https://www2.census.gov/about/partners/cac/sac/meetings/2022-03/dhc-attachment-1-safetab-dp-algorithms.pdf.

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

opaque. It is attractive to use an algorithm because it provides unambiguous documentation of the process, and is reproducible in that posted data are reproduced if the same random number seeds are used followed by the same post-DP processing. Reproducibility is important but does not imply that DP is the best approach to protect confidentiality and communicate useful information from a census or survey. It may well be that use of DP is the best strategic approach, but the case for it has not been made.

F.2 REFERENCE TABLES ON 2020 CENSUS DATA PRODUCTS AND 2010 DEMONSTRATION DATA PRODUCTS

This appendix contains two tables with details of the 2020 Census data products and implementation of the 2020 Disclosure Avoidance System (DAS) to protect the confidentiality of the data. Table F.1 provides details of the content, geographies, and timing for 2010 and 2020 Census data products, including the Redistricting File, Summary File 1/DHC File, Summary File 2/Detailed Demographic and Housing Characteristics (DDHC) Files A and B, Demographic Profile, and Public Use Microdata Sample. Table F.2 contains information on "demonstration" product releases (2010 products to which 2020 DAS specifications were applied) and other events in the implementation of the 2020 DAS during 2019–2023.

F.3 SELECTED CENSUS DATA USES

In addition to the headline uses for the U.S. decennial census, including congressional reapportionment, legislative redistricting, enforcement of the Voting Rights Act, and federal fund allocations, there are many long-standing uses by federal, state, and local agencies that merit respect. The business sector, academia, the media, and the general public also use census data. The data benchmark the Census Bureau's annual population estimates for the nation, states, counties, minor civil divisions, and incorporated places, which have their own myriad uses. In this section, we highlight selected uses of the 2020 Redistricting File and DHC File, chosen to convey the range of users and uses, and the implications of the new 2020 Disclosure Avoidance System (DAS) for those uses.[5] Uses discussed include: the housing unit-based method for state and local population estimates, public health estimates, small populations (rural America), special areas for local planning, state funding allocations to local governments, and tenure as a socioeconomic indicator.

[5]The 2020 DAS for the Redistricting and DHC Files used a TopDown Algorithm (TDA, see Chapter 11), which eliminated negative numbers from the controlled injection of random noise (based on the concept of differential privacy) and made totals consistent as blocks are added up to block groups to census tracts and so on. Effects of the TDA are based on "demonstration" files of 2010 data products to which the TDA is applied.

Copyright National Academy of Sciences. All rights reserved.

Table F.1 Census Data Products, Content, Geographies, and Release Schedule, 2010–2020

Data Product	2010 Census	2020 Census
Public Law 94-171 (P.L. 94-171) Redistricting File		
Content	5 tables—Total population by race (up to 6 or more races); Hispanic population and non-Hispanic population by race; same two tables for population ages 18 and older; housing units by occupancy status	6 tables—Same as 2010 with the addition of GQ population by major type (nursing home, dormitory, etc.)
Geographies	Nation, states, counties, all governmental units, all statistical units down to the block (smallest level of census geography)	Same as 2010
Release Schedule	12 months after Census Day (April 1), released February–March 2011	Delayed to August 12, 2021, due to COVID-19 pandemic effects on data collection
2010—Summary File 1 (SF1); 2020—Demographic and Housing Characteristics (DHC) File		
Content	Tables (including those in the Redistricting File) for persons, households, families, housing units by age, race, ethnicity, sex, household relationship, household type, housing occupancy/tenure, GQ, item imputation; cross-tabulations, including sex by age, average household size by age; iterations of selected tables by race/ethnicity of person or household/family head	Similar content to 2010 **Block-level:** Substantial deletions/moves to higher-level geography (e.g., 11 imputation rate tables deleted; 6 household/family tables deleted; 6 household/family tables moved to census tracts [often without race/ethnicity iterations]; 8 household/family/occupied housing join tables moved to S-DHC [6 for census tracts, 2 for block groups without iteration]) **Census tract-level:** Additions from block level, deletions (e.g., 11 detailed race/ethnicity tables deleted, overlapped with DDHC-A)

Copyright National Academy of Sciences. All rights reserved.

Table F.1 (cont.)

Data Product	2010 Census	2020 Census
2010—Summary File 1 (SF1); 2020—Demographic and Housing Characteristics (DHC) File (cont.)		
Geographies	**Block** as lowest level: 73 total tables = 24 person (7 iterated by race/ethnicity), 2 GQ, 13 household (3 iterated), 4 family (3 iterated), 8 person item nonresponse indicators (e.g., percent race imputed), 15 occupied units (4 iterated), 4 total or vacant housing, 3 housing item nonresponse indicators **Census tract** as lowest level: 28 total tables = 17 person (3 iterated), 3 household (1 iterated), 1 nonfamily household, 3 GQ (2 iterated), 4 occupied units **County** as lowest level: 10 total tables (GQ population)	**Block** as lowest level: 36 total tables = 15 person (2 iterated by race/ethnicity), 1 GQ, 5 household (1 iterated), 11 occupied units (2 iterated), 4 total or vacant housing **Census tract** as lowest level: 19 total tables = 5 person (4 iterated), 7 household (1 iterated), 1 family (iterated), 1 nonfamily household, 2 GQ (1 iterated), 3 occupied units **County** as lowest level: 10 total tables (GQ population) Note: Race/ethnicity iterations in some cases have more categories than 2010
Release Schedule	June–August 2011	May 25, 2023
Household/Family (Join) Tables: 2010—Part of SF1; 2020—Supplemental DHC (S-DHC) File		
Content	See SF1 above	8 total join tables = 3 household/3 family/2 occupied units tables (6 iterated by race/ethnicity) (e.g., average household size by age, tenure)
Geographies	**Block** as lowest level	**State** as lowest level (a major change announced May 31, 2023—original plans were to provide some tables for block groups, census tracts, places, and AIANNH areas)
Release Schedule	June–August 2011 (part of SF1)	Scheduled for September 2024

Copyright National Academy of Sciences. All rights reserved.

Table F.1 (cont.)

Data Product	2010 Census	2020 Census
2010—Summary File 2 (SF2) and AIAN Summary File; 2020—Detailed Demographic and Housing Characteristics (DDHC) File		
Content	Same tables as DDHC-A, B, plus cross-tabulations of households/families/occupied units by type/tenure by various characteristics of members (e.g., presence of children or nonrelatives); GQ population by type/sex/age; item imputation rates SF2—Tables iterated for 331 population groups with 100+ people in an area: total population provided for 75 race groups, 114 AIAN groups, 47 Asian groups, 43 NHOPI groups, 51 Hispanic/non-Hispanic groups AIAN File—Tables iterated for ≈ 1,600 tribes with 100+ people in an area	DDHC-A—2 tables (no tabulations for groups with < 22 people in an area): (1) total population; (2) sex by selected age categories (0, 4, 9, or 23 age categories, depending on 2020 population size) (no sex by age data for detailed groups with < 50 people nationally in 2010). Tables iterated for 300 detailed race/ethnicity groups and 1,187 AIAN tribes/villages (there are also [larger] regional groups with their own thresholds for age detail). DDHC-B—2 tables: (1) household type (depending on 2020 household estimate, total households, 2, 6, or 8 household types, depending on 2020 household estimate); (2) tenure. Tables repeated for around 300 detailed race/ethnicity groups and around 1,200 AIANNH tribes/villages
Geographies	**Census tract** as lowest level: 61 total tables = 47 person tables, 14 housing tables **County** as lowest level: 10 total tables (GQ)	DDHC-A and DDHC-B—Census tract, plus place, county, state, AIANNH areas, nation; 4 total tables
Release Schedule	SF2: December 2011–April 2012 AIANSF: December 2012	DDHC-A—September 21, 2023 DDHC-B—Scheduled for September 2024

Copyright National Academy of Sciences. All rights reserved.

Table F.1 (cont.)

Data Product	2010 Census	2020 Census
Demographic Profile		
	22 person and housing indicators (e.g., median age by sex); **Census tract** as lowest level; Released May 2011	18 indicators (missing are household-size variables, to be included in S-DHC); **Census tract** as lowest level; Released with DHC in May 2023
Public Use Microdata Sample (PUMS) File		
	Hierarchical household-person file containing all variables in the census for geographies of about 100,000 people (public use microdata areas, or PUMAs) for 10% of the population Released near end of 2014	Plans announced March 27, 2023, to produce separate privacy-protected microdata files for 100% of persons and housing units with the Redistricting and DHC variables for blocks; release date not yet scheduled

NOTES: AIAN, American Indian or Alaska Native; AIANNH, American Indian/Alaska Native/Native Hawaiian; GQ, group quarters; NHOPI, Native Hawaiian or Other Pacific Islander. For "Block (or census tract) as lowest level," all higher-level geographies are included (e.g., block group, census tract, incorporated place, minor civil division, school district, county, AIANNH area, state).

SOURCES: U.S. Census Bureau (2012a,b, 2022a, 2023a,b); Devine et al. (2023).

Copyright National Academy of Sciences. All rights reserved.

Table F.2 Demonstration Product Releases and Other Events in the Implementation of the 2020 Disclosure Avoidance System, 2019–2023

Test/Event/Date	Description	Comments
Spring 2019: End-to-End Census Test Redistricting File Prototype	First use of differential privacy-based TopDown Algorithm (TDA) to protect redistricting data from 2018 Rhode Island census test; invariants—county population, block level total/occupied housing units, group quarters facilities by type; $\epsilon = 0.25$	Census Bureau informed users that the product was to test systems and code; very low ϵ meant data were not useful (Fontenot, 2019)
Spring 2019: 1940 Census Redistricting Data Comparison	TDA used for redistricting data from the 1940 Census, using enumeration districts as the lowest geographic level; ϵ ranging from 0.25–8 (actual 1940 data are public)	See Asquith et al. (2019) and Brummet et al. (2019) for evaluations of the utility and privacy protection of the 1940 data
October 2019: 2010 Demonstration Product Baseline (detailed summary metrics released March 2020)	TDA used for redistricting and selected DHC tables using 2010 data as demonstration; $\epsilon = 4$ (person tables), $\epsilon = 2$ (housing tables)	Basis for CNSTAT December 2019 Workshop; see U.S. Census Bureau (2019) and associated summary metrics under "2010 Demonstration Data Products Baseline 2019-10-29" at https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-development.html
December 2019: CNSTAT Workshop on 2010 Demonstration Baseline	Proceedings: National Academies of Sciences, Engineering, and Medicine (2020); presentations and video at: https://www.nationalacademies.org/event/12-11-2019/workshop-on-2020-census-data-products-data-needs-and-privacy-considerations	Presentations by many users revealed substantial inaccuracies for small off-spine geographies and small population groups

Copyright National Academy of Sciences. All rights reserved.

Table F.2 (cont.)

Test/Event/Date	Description	Comments
May 2020: Demonstration Product Update #1 (summary metrics released with this and later updates)	TDA used for redistricting and DHC tables; $\epsilon = 4$ (person tables) and $\epsilon = 2$ (housing tables) but update file did not include housing tables due to time constraints; *improvement:* optimize the post-processing steps	Fact sheet and associated summary metrics available under 2020-05-27 run at https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-development.html
September 2020: 2010 Demonstration Product Update #2; November 2020: File Correction	TDA used for redistricting tables only; $\epsilon = 4$ (person tables); $\epsilon = 0.5$ (housing table); *improvements:* put AIAN tribal areas in their own geographic hierarchy within each state; add state-level population invariant for total AIAN population	Fact sheet and associated summary metrics available under 2020-09-17 run at https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-development.html
April 2021: 2010 Demonstration Product Update #3	TDA used for redistricting tables; two sets of files and metrics, one with same ϵ as Update #2, other with $\epsilon = 10$ (person tables), $\epsilon = 2$ (housing table); *improvements:* replace discrete Laplace noise infusion with zero-concentrated differential privacy (zCDP) discrete Gaussian mechanism to reduce outliers; use accuracy target for largest racial-ethnic group in any geography < 500 people of ±5 percentage points of enumerated value 95% of the time; optimize to bring off-spine geographies closer to the spine; separate post-processing for group quarters at the block group level	In zCDP, ρ and δ replace (and can be converted to) ϵ as the privacy-accuracy parameters; fact sheet and associated summary metrics available under 2021-04-28 run at https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-development.html
June 2021: 2010 Demonstration Product Update #4/ Redistricting File Production Settings	TDA used for redistricting tables; $\epsilon = 17$ (person tables), $\epsilon = 2.5$ (housing tables); *improvement:* allocate more PLB to total population and race/ethnicity for block groups and above	Fact sheet and associated summary metrics available under 2021-06-08 run and 2021-08-12 production of Redistricting File at https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-development.html

Copyright National Academy of Sciences. All rights reserved.

Table F.2 (cont.)

Test/Event/Date	Description	Comments
March 2022: Demonstration Product Update #5 for DHC File	TDA used for DHC tables (person tables released ahead of housing tables); for consistency, DHC inherited person/housing redistricting PLB ($\epsilon = 19.61$) and added PLB for combined DHC ϵ (person and housing) of 39.91; *improvement:* incorporate building block geographies for population estimates into spine	Fact sheet and associated summary metrics available under 2022-03-16 run at https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-development.html
June 2022: CNSTAT Workshop on DHC Demonstration Product Update #5	National Academies of Sciences, Engineering, and Medicine (2023); presentations and video at: https://www.nationalacademies.org/event/06-21-2022/2020-census-data-products-workshop-on-the-demographic-and-housing-characteristics-files	Presentations identified usable data but also population groups/areas with less accuracy in DHC demo product (e.g., AIAN population, Liebler; renters, Reynolds and Vink; rural population , Mueller; denominators for local age-adjusted death/disease rates, Werner; small Traffic Analysis Zones, Kaneff)
August 2022: 2010 Demonstration Product Update #6/ DHC Production Settings	TDA used for DHC tables; combined DHC $\epsilon = 46$; *improvements:* changed lowest level of geography for some tables from state/county to census tract; iterated tables for sex by single year of age at census tract level; greater accuracy for some tables but not group quarters people by age; limited reduction in person-housing unit inconsistencies	See https://www2.census.gov/programs-surveys/decennial/2020/program-management/data-product-planning/2010-demonstration-data-products/02-Demographic_and_Housing_Characteristics/2022-08-25_Summary_File/2022-08-25_Factsheet.pdf; Census Data Stewardship Executive Policy Committee decided final PLB parameters for DHC, November 2022 (made public in April 2023; see below); DHC released May 25, 2023
January 2023: 2010 "Proof of Concept" Product for DDHC-A File[a]	SafeTab-P algorithm used for DDHC-A tables; consistency not promised internally or with other products; negatives possible; all margins of error determined in advance and met 95% of the time; adaptive design using thresholds to determine age detail; no data published for negatives or "alone or in combination"; detailed race/ethnicity groups < 22 people at substate level	See https://www2.census.gov/programs-surveys/decennial/2020/program-management/data-product-planning/2010-demonstration-data-products/03-Detailed_DHC-A/2023-01-31/FAQs-Detailed_DHC-A.pdf; "Proof of Concept" includes summary metrics of accuracy and not tables or microdata; users cautioned against custom aggregations; DDHC-A released September 21, 2023

Copyright National Academy of Sciences. All rights reserved.

Table F.2 (cont.)

Test/Event/Date	Description	Comments
April 2023: 2010 Demonstration Data Products Suite—Redistricting and DHC	TDA used for 2010 Demonstration PPMF combining the differentially privately protected redistricting and DHC variables with final 2020 production ϵ values (combined person and housing, redistricting, and DHC, $\epsilon = 53$); plus separate "noisy measurement files," with noise infused but before running the TDA for consistency and non-negativity;[b] plus summary metrics, which show accuracy improvements for householder race, presence/age of own children, relationship to householder; same-sex married and unmarried partners; single years of age for children for counties and school districts; age for persons in group quarters; same-sex married and unmarried partners; accuracy decreased for 5-year age groups.	See https://www.census.gov/data/academy/webinars/2023/2010-demonstration-data-products-suite.html; release responds to researcher requests to see the separate effects of noise infusion and the TDA on differences between the 2010 data protected with the 2020 DAS and the original 2010 data (e.g. Groshen and Goroff, 2022). Note: 2010 Demonstration PPMF variables, reflecting 2020 definitions, not always consistent with 2010 Summary File 1 variables

[a] 2010 "Proof of Concept" product for DDHC-B File scheduled for release January 2024; no Proof-of-Concept product planned for Supplemental-DHC.
[b] Noisy Measurement Files released for the 2020 Redistricting and DHC Files in June 2023.

NOTES: TDA, Top Down Algorithm; DHC, Demographic and Housing Characteristics (File); DDHC, Detailed DHC; CNSTAT, Committee on National Statistics of the National Academies of Sciences, Engineering, and Medicine; PLB: privacy loss budget; Updates #2–4 released as privacy protected microdata files (PPMFs); updates #5–6 released as tables; IPUMS processed both PPMFs and tables to be usable by others: Van Riper et al. (2023).

SOURCES: See table entries.

Copyright National Academy of Sciences. All rights reserved.

F.3.1 Housing Unit-Based Method for State and Local Population Estimates

Many states produce their own postcensal population estimates, often using a housing unit-based method. This method starts with census data on housing, occupied housing, and persons per household, applied to local data on building permits. In turn, local population estimates are used for myriad purposes. For 2020, however, estimates of occupied housing and household size derived from the Redistricting File are highly variable, particularly for small geographies. The reason is that the TopDown Algorithm (TDA) processes people and housing units separately.

Hardcastle (2019) described a Nevada revenue-sharing program, begun in 1987, which currently allocates about $79 million among 82 governmental units using state-generated population estimates, which average a housing unit-based estimate and a regression-based estimate. About 20% of the Nevada governmental units that are eligible for state revenue sharing had fewer than about 400 people in 2020[6] and thus would likely have highly variable population estimates from the housing unit-based method using data from the 2020 Redistricting File. More accurate estimates of persons per occupied housing units are now available in the DHC File released in May 2023.[7]

F.3.2 Public Health Estimates

Public health analysts require small-area census data for planning, implementation, and evaluation of public health practices. Commonly needed rates include: (1) crude rates of the overall average rate of a disease, such as the total number of cancer cases divided by the total population; (2) age- or sex-specific rates, such as the number of cases for an age group (e.g., 25–34-year-olds) divided by the population of the age group; and (3) age-standardized rates, which allow for comparing health outcomes among population groups with differing age distributions—these rates apply a standard age distribution to the populations being compared.

Werner (2022) described typical public health applications of census data and the effects that the TDA could have on them by comparing the 2010 DHC Demonstration File (March 2022 version) to the original 2010 Summary File 1 (SF1). For age-adjusted estimates of COVID-19 rates by county and race, there were substantial percentage differences in rates for groups other than White people, particularly in nonmetropolitan counties. Such differences

[6]From the 2020 Redistricting File—see the state tables at https://mdi.georgetown.edu/census-geographies-project/.

[7]The DHC File has more accurate estimates because a person per occupied unit recode for each household head was added to the tables to which controlled random noise and the TDA were applied, whereas all person and housing variables were processed separately in the Redistricting File.

Copyright National Academy of Sciences. All rights reserved.

could mean that interventions or scarce medical resources would be targeted incorrectly. Overall age-standardized county-level heart disease death rates were reasonably comparable between the two data sets, but rates for age and sex groups often showed substantial differences for smaller populations and areas (under 10,000 people). For age-adjusted estimates of hospitalizations and emergency department visits for asthma, county data were reasonably comparable between the 2010 DHC Demonstration File and the original 2010 SF1, but much less so for census tracts. Block data by race, ethnicity, age, and sex to assess populations living near environmental hazards proved problematic for characterizing risks accurately and targeting populations at higher risk. Similar analyses of the 2010 DHC Demonstration File (August 2022 version) compared with the original 2010 SF1 found only "minimal improvements."[8]

F.3.3 Small Populations—Rural America

Mueller (2022) discussed the implications of the TDA for rural areas. He said that rural America lost population on net for the first time between 2010 and 2020 and is prone to high rates of poverty and poor health. He said these trends are well known but hard to document because many datasets have much-too-high variance in estimates for rural areas (e.g., estimates for census tracts from the American Community Survey are unusable for rural areas). The census has long been the gold standard for rural demography, but even then, demographers use the county as the unit of analysis.

Mueller analyzed population and median age by race and ethnicity for counties classified by Rural Urban Continuum Code from code 1 (metropolitan county of 1 million or more people) to code 9 (county not adjacent to a metropolitan area with fewer than 2,500 people). Comparing the 2010 DHC Demonstration File (March 2022) and the original 2010 SF1, Mueller found large percentages of highly rural counties (codes 7–9) with differences of 10% or more in total population for non-Hispanic Black people, Hispanic people, and non-Hispanic American Indian and Alaska Native (AIAN) people.[9] He concluded:

> It is not clear to the rural demography community how differential privacy makes sense when accuracy for small groups is important. … Prior methods seemingly did a much better job of preserving headcount totals. Small groups and places have just as much of a right to accuracy and representation as the large groups and places. I know many have

[8]Document sent from B. Moyer, National Center for Health Statistics, to R. Jarmin, U.S. Census Bureau, attaching the updated U.S. Centers for Disease Control and Prevention response; provided by A. Werner to K. Stone, September 29, 2022.

[9]Total population in the DHC File is the same as in the Redistricting File, so the fact that the DHC Demonstration File that Mueller used for his analysis does not reflect the final production values for epsilon is not relevant.

Copyright National Academy of Sciences. All rights reserved.

> concerns about reidentification, but if our 2020 data is released with these discrepancies, I am not sure I will be able to use it or recommend others use it. Sadly, we don't have an alternative source of data for rural areas in the United States.

F.3.4 Special Areas for Local Planning

State and local governments and regional planning organizations often aggregate census blocks, block groups, or census tracts to form their own areas for planning purposes. Kaneff (2022) described the myriad uses of data for traffic analysis zones (TAZs), built up from blocks, by the Northern Virginia Regional Commission and the local governments in the region—for example, siting fire stations, schools, parks, and other local facilities; transportation planning, including projections; emergency planning; and other uses. Of 1,415 TAZs, 503 had fewer than 450 people—household size data for these areas from the 2020 Redistricting File were not fit for use in the Commission's judgment. Data for TAZs with 450–1,000 people were better, and data for TAZs with 1,000 or more people were reliable, based on the 2010 demonstration data comparisons.

Alvarez and Maurer (2022) described the demographic uses of New York City's approximately 200 Neighborhood Tabulation Areas (NTAs). Their analysis found that the March 2022 version of the 2010 DHC Demonstration File, compared with the original 2010 SF1, was fit for use for examining 5-year age distributions for NTAs. This finding is not surprising in that each NTA on average consists of about 10 census tracts totaling about 40,000 people.

F.3.5 State Funding Allocations to Local Governments

Many states allocate their own funds to their local governments. For example, Craigle (2022) identified several state programs in Montana that allocate funding to counties. Hardcastle (2019) described a long-standing Nevada state revenue-sharing program for local governments. Nagle and Kuhn (2019) provided an example of a specific Tennessee statute (Title 67) that allocates funds to incorporated places. The statute reads:

> (3)(A) Four and six thousand thirty ten-thousandths percent (4.6030%) [of state sales and uses taxes] shall be appropriated to the several incorporated municipalities within the state of Tennessee to be allocated and distributed to them monthly by the commissioner of finance and administration, in the proportion as the population of each municipality bears to the aggregate population of all municipalities within the state, according to the latest federal census and other censuses authorized by law.

Continuing with the Tennessee example, as of 2020, the state had 344 functioning incorporated places, ranging in population size from 66–633,000

Copyright National Academy of Sciences. All rights reserved.

people, with a median size of about 2,000 people.[10] Twenty percent of Tennessee incorporated places had fewer than 550 people. Each person represented about $115 in allocated funds from Tennessee state sales and use tax revenue. Population totals in the Redistricting File, based on 2010 demonstration data, are quite accurate overall. About 5% of incorporated places with 500 or fewer people in the nation, however, have errors of 5% or more (25 people or more) and 11% have errors of 2–5% (10–25 people), comparing the privacy-protected 2010 Redistricting File with the publicly released 2010 Redistricting File.[11] Consequently, among Tennessee's about 70 incorporated places with fewer than 500 people, assuming the national error rates apply, 4 places could lose and 4 could gain 25 or more people due to noise infusion, and another 8 places could lose and 8 could gain, say, 20 people. In dollar terms, the losses (gains) would be about $2,000–3,000 or more. These are not large differences, but they could matter for the budget of a small incorporated place. Nagle and Kuhn (2019), based on their analysis, assert:

> Protecting characteristics is admirable, but the number of people in a place has never been assumed private. State constitutions assume that Census population counts for municipalities are accurate. The starting point for DP [differential privacy] should have been change in log probability conditional on population count being public information.

F.3.6 Tenure as a Socioeconomic Indicator

Tenure—whether a household owns or rents its residence—is the one socioeconomic variable collected in the census. Homeownership rates are an important indicator of wealth inequity among areas and among race and ethnic groups. It is a variable of interest for government policy and for research, including analysis of racial disparities and geographic segregation. Reynolds and Vink (2022) compared the March 2022 version of the 2010 DHC Demonstration File with the original 2010 SF1 and found that the two data sets differed more for households in rental-majority areas (census tracts) compared with owner-majority areas and that the differences were particularly pronounced for households with children and large households in rental-majority areas. Differences were also large for aggregations of census tracts to substate levels.

Replicating the analysis with the August 2022 version of the 2010 DHC Demonstration File, Reynolds and Vink (2023) found some improvements but also that differences between majority owner and renter areas remained. Average homeownership rates by race and ethnicity varied significantly between

[10] From the 2020 Census Redistricting File—see the state tables at https://mdi.georgetown.edu/census-geographies-project/.

[11] See https://www2.census.gov/programs-surveys/decennial/2020/program-management/data-product-planning/2010-demonstration-data-products/.

Copyright National Academy of Sciences. All rights reserved.

the demonstration and original datasets for most race and ethnic groups, with the exceptions of Asian people and people who were not Hispanic.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.

– G –

Biographical Sketches of Panel Members and Staff

TERESA A. SULLIVAN *(chair)* is University Professor of sociology and president emerita of the University of Virginia. President of the University of Virginia from 2010–2018, she previously served as provost at the University of Michigan, executive vice chancellor for academic affairs of the University of Texas system, and numerous administrative positions (including vice president and graduate dean) at the University of Texas at Austin. In 2019–2020, she was named interim provost of Michigan State University. A demographer with extensive background in census data, she authored *Census 2020: Understanding the Issues* (2020, Springer) and delivered the president's invited address, "Coming to our Census: How Social Statistics Underpin Our Democracy (And Republic)," at the 2019 Joint Statistical Meetings in Denver. She is past chair and member of the U.S. Census Bureau's Advisory Committee on Population Statistics. At the National Academies of Sciences, Engineering, and Medicine (the National Academies), she chaired the Panel on Higher Education Productivity and co-chaired the Committee on Improving Higher Education's Responsiveness to STEM Workforce Needs, and she has served on several other panels including the Panel on Immigration Statistics and the Panel on the Requirements for the Year 2000 Census and Beyond. She is a fellow of the American Academy of Arts and Sciences and the American Association for the Advancement of Science. She has a B.A. from Michigan State University and A.M. and Ph.D. degrees in sociology from the University of Chicago.

Copyright National Academy of Sciences. All rights reserved.

MARGO ANDERSON is Distinguished Professor, emerita, of history and urban studies at the University of Wisconsin–Milwaukee. Her research specializes in the history of the social sciences and the development of statistical data systems and in urban history. She is the author of *The American Census: A Social History* (first edition 1988 and second edition 2015, Yale University Press), lead editor of the first edition of the *Encyclopedia of the U.S. Census* (2000, CQ Press) and co-editor of the second edition (2011, CQ Press), and co-author with Stephen E. Fienberg of *Who Counts? The Politics of Census-Taking in Contemporary America* (1999, Russell Sage; revised edition, 2001). She is past president of the Social Science History Association and a fellow of the American Statistical Association, and she has chaired the council of the Interuniversity Consortium for Political and Social Research. At the National Academies, she served as a member of the Panel on Census Requirements in the Year 2000 and Beyond. She has a B.A. from Bucknell University and M.A. and Ph.D. degrees from Rutgers University, all in history.

ROBERT M. BELL retired as statistical researcher at Google in 2017. Previously, he held extensive tenures at AT&T Labs–Research and at RAND. His research interests include survey research methods and machine learning methods, record linkage, and statistical issues in a wide range of public policy applications. His team won the $1 million Netflix Prize in 2009, a competition to build algorithms to predict Netflix customer movie ratings most accurately. At the National Academies, he is a past member of the Committee on National Statistics (CNSTAT) and the Division Committee for the Behavioral and Social Sciences and Education. He chaired the Panel on Correlation Bias and Coverage Measurement in the 2010 Census and was a member of the Panel to Review the 2000 Census, and he has been a member of several other National Academies panels. He is a fellow of the American Statistical Association. He has a B.S. degree in mathematics from Harvey Mudd College, an M.S. degree in statistics from the University of Chicago, and a Ph.D. in statistics from Stanford University.

CONSTANCE F. CITRO is a senior scholar in the Committee on National Statistics (CNSTAT) at the National Academies. She previously served as CNSTAT director and senior study director. Prior to joining CNSTAT in 1984, she held positions as vice president of Mathematica Policy Research, Inc. and Data Use and Access Laboratories, Inc. She is a fellow of the American Statistical Association and an elected member of the International Statistical Institute. She received the Roger Herriot Award for Innovation in Federal Statistics in 1997 and the Waksberg Award in Survey Methodology in 2014. She has a B.A. in political science from the University of Rochester and an M.A. and Ph.D. in political science from Yale University. For CNSTAT, she directed evaluations of decennial censuses, the Survey of Income and Program Participation, microsimulation models for social welfare programs,

Copyright National Academy of Sciences. All rights reserved.

estimates of school-age poverty for small geographic areas, and the study that produced *Measuring Poverty: A New Approach* (1995)—the genesis of the Supplemental Poverty Measure (SPM). She assisted the study that produced *A Roadmap to Reducing Child Poverty* (2019) and is currently assisting studies on the 2020 Census, improvements to the SPM, and an integrated system of household income, consumption, and wealth statistics. She has edited or co-edited the 2nd–7th editions of CNSTAT's flagship publication, *Principles and Practices for a Federal Statistical Agency*.

MICHAEL L. COHEN is a senior program officer for the Committee on National Statistics at the National Academies. Previously, he was a mathematical statistician at the Energy Information Administration and held positions at the School of Public Affairs at the University of Maryland and at Princeton University. His general area of interest is the use of statistics in public policy, with particular focus in census undercount, model validation, and robust estimation. He is a fellow of the American Statistical Association and an elected member of the International Statistical Institute. He has a B.S. in mathematics from the University of Michigan and an M.S. and a Ph.D. in statistics from Stanford University.

DANIEL L. CORK *(study director)* is a senior program officer for the Committee on National Statistics (CNSTAT) at the National Academies. He joined the CNSTAT staff in 2000 and has served as study director or program officer for almost all subsequent census- or American Community Survey-related studies, including the Panels on Residence Rules in the Decennial Census, Research on Future Census Methods (2010 Planning panel), Review of the 2000 Census, and Review of the 2010 Census, and the Standing Committee on Reengineering Census Methods. He also directed the Panel to Review the Programs of the Bureau of Justice Statistics and the Panel on Modernizing the Nation's Crime Statistics (in cooperation with the Committee on Law and Justice), was senior program officer for the Panel on the Feasibility, Accuracy, and Technical Capability of a National Ballistics Database (joint with the Committee on Law and Justice and the National Materials Advisory Board), and contributed to the work of the Committee on Best Practices for Assessing Mortality and Significant Morbidity Following Large-Scale Disasters (Board on Health Sciences Policy). His research interests include quantitative criminology, geographical analysis, Bayesian statistics, and statistics in sports. He has a B.S. in statistics from George Washington University and an M.S. in statistics and a joint Ph.D. in statistics and public policy from Carnegie Mellon University.

KATHRYN EDIN is William Church Osborn Professor of sociology and public affairs at Princeton University. Previously, she was on the faculty of the Institute for Policy Research at Northwestern University and an associate fellow of the Institute for Research on Poverty at the University of Wisconsin/Madison and of the Joint Center for Poverty Research at

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

Northwestern University/University of Chicago. Her principal research interests include poverty and social inequality, urban and community sociology, and family and gender. She is currently principal investigator or co-principal investigator on research projects on the social role of children and marriage among low-income adults, the role of local labor markets and child support programs in affecting fathers' economic and emotional involvement with their children, and couple dynamics among low-income married and unmarried couples with young children. Previously, she held faculty or research posts at University of Pennsylvania, Rutgers University, and the Russell Sage Foundation. Elected a member of the National Academy of Sciences in 2014, she served on the Panel on Residence Rules in the Decennial Census and Panel on Data and Methods for Measuring the Effects of Changes in Social Welfare Programs. She has a B.A. in sociology from North Park University and M.A. and Ph.D. degrees in sociology from Northwestern University.

Marc Hamel retired as Director General of the Census Management Office at Statistics Canada in March 2020, after a 34-year career with the national statistical office. His management of the 2011 and 2016 quinquennial Canadian censuses marked the successful expansion of internet response to census inquiries and of related innovative methodologies, as well as major improvements to census operational systems. In this position, he also initiated and directed a large research program on the use of administrative sources to replace direct enumeration in future censuses. During his tenure on the Census Program, he led international groups of census experts in the redevelopment of the Principles and Recommendations for Censuses of Population and Housing under the United Nations Statistical Division (2015) and those under the United Nations Economic Commission for Europe (UNECE; 2014) in preparation for the 2020 international round of censuses. He also was the chair of the Steering Group of census experts under UNECE from 2010 to 2019, and a member of the International Census Forum (Canada, United States, United Kingdom, Ireland, Australia, and New Zealand). Since his retirement, he has written a paper on the impacts of the COVID-19 pandemic on censuses (commissioned by the UNECE), and was a consultant with the UK Office for National Statistics (on contingency planning for their 2021 census) and with Statistics Canada (on the changing risk posture for the 2021 Canadian census). He has a bachelor's degree in social sciences (sociology and criminology) from Ottawa University.

George T. Ligler is professor and Dean's Excellence Chair of Multidisciplinary Engineering at Texas A&M University. He is also the proprietor of GTL Associates, a consultancy which has provided interdisciplinary systems integration/engineering and product management services to over 40 clients on three continents. He has extensive experience in information

Copyright National Academy of Sciences. All rights reserved.

management and software and computer system engineering, and previously worked at Burroughs Corporation (1980–1982) and Computer Sciences Corporation (1984–1988). An elected member of the National Academy of Engineering (NAE) and former Chair of NAE Section 12, Special Fields and Interdisciplinary Engineering, he has additionally received a number of national awards in aviation. He has also participated in ten National Academies of Sciences, Engineering, and Medicine committees (chairing three of those committees) advising the Departments of Transportation, Treasury, and Commerce, and is a current member of the Academies' Aeronautics and Space Engineering Board. He began work on census issues as a member of the Panel on Research on Future Census Methods (2010 Census planning) and was tapped as a member of the expert committee separately formed by the U.S. Secretary of Commerce to advise on options for the Census Bureau's replan of its Field Data Collection Automation Contract in early 2008. He subsequently served as a member of the Panel to Review the 2010 Census and its subsequent Standing Committee on Reengineering Census Operations, those panels playing an important role in the design of the 2020 Census. A Rhodes scholar, he received his B.S. in mathematics from Furman University in 1971 and his M.Sc. and D.Phil. from Oxford University.

THOMAS A. LOUIS is professor emeritus of biostatistics at the Johns Hopkins Bloomberg School of Public Health. From 2018–2020 he was an expert statistical consultant to the Food and Drug Administration's Center for Drug Evaluation and Research and from 2013–2015, he served as associate director for research and methodology and chief scientist of the U.S. Census Bureau; each of these via interagency personnel agreements with Johns Hopkins. Previously, he was on the faculties of Boston University, Harvard School of Public Health, and University of Minnesota School of Public Health, and was senior statistical scientist at RAND. He is a fellow of the American Statistical Association, American Association for the Advancement of Science, and the Institute of Mathematical Statistics as well as an elected member of the International Statistical Institute. At the National Academies, he chaired the Panel on a Review of Statistical Issues in the Allocation of Federal and State Program Funds and is a past member of both the Committee on National Statistics and the Committee on Applied and Theoretical Statistics, among other National Academies' panel memberships, and is a National Associate. He has a B.A. in mathematics from Dartmouth College and a Ph.D. in mathematical statistics from Columbia University.

LLOYD B. POTTER has been state demographer of Texas since May 2010. In that capacity, he is director of the Texas Demographic Center at University of Texas at San Antonio (UTSA), where he is also professor of demography. Prior to joining the UTSA faculty and becoming state demographer, he

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

directed the Center for Study and Prevention of Injury, Violence, and Suicide at Education Development Center, Inc., and the youth violence and suicide prevention program at the U.S. Centers for Disease Control and Prevention. His work focuses on application of demographic and social science methods in public health settings. He has a B.S. in sociology from Texas A&M University, an M.S. in education from the University of Houston at Clear Lake, an M.P.H. from Emory University, and a Ph.D. in demography and sociology from the University of Texas at Austin.

JOSEPH J. SALVO retired as chief demographer, Population Division, New York City Department of City Planning, in spring 2021, having previously served as director of the division. He is now affiliated with the National Conference on Citizenship and is institute fellow in the Social and Decision Analytics Division, University of Virginia Biocomplexity Institute. He has been a member of the Committee on Population Statistics of the Population Association of America and served as past president of the Association of Public Data Users. He was one of the data analysts engaged by the American Statistical Association's task force on census quality indicators. For the National Academies, he co-chaired the planning committee for the Workshop on 2020 Census Data Products and has continued as effective co-chair of the series of follow-up expert meetings on disclosure avoidance in the 2020 Census. He was a member of the Panel on Research on Future Census Methods (2010 Census planning) and of three panels related to the Census Bureau's American Community Survey, and co-chaired the planning committee for the Workshop on Respondent Burden in the American Community Survey. Dr. Salvo is a fellow of the American Statistical Association. He has M.A. and Ph.D. degrees in sociology from Fordham University.

REGINA SHIH is professor of epidemiology at the Rollins School of Public Health, Emory University, as of April 2023. She joined the Emory faculty after serving as senior policy researcher and director of the Social and Behavioral Policy program at RAND, where she retains an affiliated adjunct appointment. Previously, she headed the Climate Change and Health group at RAND. Her research focuses on environmental health and behavioral health, and her projects have included developing a national database for assessing cognitive aging and dementia risk based on sociodemographic and physical characteristics of neighborhoods. Her work involves mixed-method studies involving large epidemiological study designs and statistical analyses such as latent variable modeling. She serves on the board of the National Alliance for Caregiving. She has a B.A. in neuroscience from Johns Hopkins University and a Ph.D. in psychiatric epidemiology from the Johns Hopkins Bloomberg School of Public Health.

C. MATTHEW SNIPP is Burnet C. and Milfred Finley Wohlford professor of sociology at Stanford University. At Stanford, he has been director of Secure

Copyright National Academy of Sciences. All rights reserved.

Data Center, deputy director of the Institute for Research in the Social Sciences, chair of the Native American Studies program, and director of the Center for Comparative Studies of Race and Ethnicity. He has written extensively on American Indians, focusing specifically on the interaction of American Indians and the U.S. census. Prior to moving to Stanford, he was associate professor and professor of rural sociology at the University of Wisconsin–Madison, where he held affiliate appointments with several other units, and assistant and associate professor of sociology at the University of Maryland. A current member of the Committee on National Statistics, he has served on numerous census-related panels for the National Academies, including the Panel on Research on Future Census Methods (2010 planning), Panel on Residence Rules in the Decennial Census, Panel to Review the 2010 Census, and the Standing Committee on Reengineering Census Operations. He has also served on the Census Bureau's Racial and Ethnic Advisory Committee and on the Board of Scientific Counselors for the U.S. Centers for Disease Control and Prevention/National Center for Health Statistics. He has a M.S. and Ph.D. in sociology from the University of Wisconsin, Madison.

KATRINA BAUM STONE is a senior program officer for the Committee on National Statistics at the National Academies. She was one of the first researchers at the Cartographic Modeling Lab and has since worked with multiple federal agencies as a contractor or public servant including as government program director for J.D. Power, an expert appointment at the U.S. Peace Corps, and more than a decade at the U.S. Department of Justice as a senior statistician in the Bureau of Justice Statistics and senior research officer in the National Institute of Justice. Her research focuses on risk and resilience for vulnerable populations such as children, veterans, and survivors of sexual assault, stalking, and human trafficking. She served on the Federal Interagency Forum on Child and Family Statistics and as a founding member on the Institute of Medicine's Forum on Global Violence Prevention. She has a B.A. in law and society from University of California, Santa Barbara, an M.S. in criminal justice from Northeastern University, and a Ph.D. in social welfare from the University of Pennsylvania.

EDWARD TELLES is distinguished professor of sociology at the University of California, Irvine, where he is director of the Center for Research on International Migration. Prior to moving to Irvine, he was on the faculties of the University of California, Los Angeles and Santa Barbara campuses, and Princeton University. He has also been program officer in human rights at the Ford Foundation's Rio de Janeiro office, visiting scholar at the Russell Sage Foundation, and visiting professor at the University of Campinas and the Federal University of Bahia, both in Brazil. He has directed major surveys of race and ethnicity in Brazil, Colombia, Mexico, Peru, and the United States, including the Project on Ethnicity and Race in Latin America

Copyright National Academy of Sciences. All rights reserved.

Assessing the 2020 Census: Final Report

(PERLA) project and the Mexican American Study Project (MASP). He is a member of the American Academy of Arts and Sciences. He has a B.A. in anthropology from Stanford University, an M.A. in urban planning from the University of California, Los Angeles, and a Ph.D. in sociology from the University of Texas at Austin.

WENDY UNDERHILL is director of elections and redistricting at the National Conference of State Legislatures (NCSL), based in Denver, Colorado. In that capacity, she provides bipartisan information and analysis on redistricting law and processes, as well as election administration issues, and so serves as an important conduit for census and related information to the redistricting community. Prior to joining NCSL in 2010, she worked as a freelance writer, a securities analyst at Incubix, and a budget and policy analyst at the U.S. Department of the Treasury and U.S. Senate Committee on the Budget. She has a B.A. in English literature from Northwestern University and an M.P.A. from Wichita State University.

DAVID VAN RIPER is director of spatial analysis at the Minnesota Population Center at the University of Minnesota, beginning work at Minnesota as a graduate research assistant in 2001. He is co-principal investigator on two Integrated Public Use Microdata Series (IPUMS) products, the National Historical Geographic Information System (NHGIS) that provides both historical and contemporary small-area data in GIS format for the United States and GeoMarker, which provides secure geocoding and small-area data linkage services. His research interests include population health and health systems and the analysis of structural racism and health inequities. For the National Academies, he participated in the December 2019 Workshop on 2020 Census Data Products and, as participant in the subsequent expert meetings on application of differential privacy to the 2020 Census, provided essential service to census stakeholders by assuming tabulation and dissemination functions for multiple iterations of the Privacy Protected Microdata Files through the NHGIS site. He has a B.A. from the University of Wisconsin–Madison and an M.A. from the University of Minnesota, both in geography.

Copyright National Academy of Sciences. All rights reserved.

COMMITTEE ON NATIONAL STATISTICS

The Committee on National Statistics was established in 1972 at the National Academies of Sciences, Engineering, and Medicine to improve the statistical methods and information on which public policy decisions are based. The committee carries out studies, workshops, and other activities to foster better measures and fuller understanding of the economy, the environment, public health, crime, education, immigration, poverty, welfare, and other public policy issues. It also evaluates ongoing statistical programs and tracks the statistical policy and coordinating activities of the federal government, serving a unique role at the intersection of statistics and public policy. The committee's work is supported by a consortium of federal agencies through a National Science Foundation grant, a National Agricultural Statistics Service cooperative agreement, and several individual contracts.

Copyright National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.