



A NEW APPROACH FOR THE 2030 DECENNIAL CENSUS

A 21st Century Census for a 21st Century Society

February 4, 2021

Sallie Keller

Distinguished Professor in Biocomplexity,
Division Director
Social and Decision Analytics Division
<https://orcid.org/0000-0001-7303-7267>
sak9tr@virginia.edu

Stephanie Shipp

Deputy Director and Professor,
Social and Decision Analytics Division
<https://orcid.org/0000-0002-2142-2136>
sss5sc@virginia.edu

Acknowledgments: We would like to thank the Alfred P. Sloan Foundation, Grants G-2019-11316 and G-2020-14002 for their support of this research. We would also like to thank Ken Prewitt, Columbia University, John Thompson, Biocomplexity Distinguished Institute Fellow, and others on our Steering Committee for their review of this work.

Citation: Keller, S., and Shipp, S. (2022). A New Approach for the 2030 Decennial Census - A 21st Century Census for a 21st Century Society. Proceedings of the Biocomplexity Institute, TR 2022-022. <https://doi.org/10.18130/3was-x496>



**BIOCOMPLEXITY
INSTITUTE**

Abstract

The founding fathers realized a census was necessary to establish the republican form of government based on the representation of the resident population as it distributed itself across the original 13 states and then spread into new territories, soon to become new states. The decennial census was a mechanism to apportion the House of Representatives based on each state's population size. It was early recognized that the massive effort to count everyone in the growing country and locate them at their permanent residence could provide additional information inexpensively, which gradually led to an enlarged census of the social, demographic, and housing characteristics of the total U.S. population.

To meet the country's needs, we propose a new process whereby essential data will be built over the decade to provide a complete and accurate enumeration of the population of the United States for the first time in 2030. A vital component of this enumeration will be to meet the constitutional requirements of the 2030 Census to provide state population totals. This enumeration will also be used as a benchmark for tabulations created throughout the decade by the augmentation of the 2020 decennial census with multi-source federal and private sector sources. We call this an *Essential-Data Count of the Resident Population* coupled with a *21st Century Census Information Platform*.

A New Approach for the 2030 Decennial Census - A 21st Century Census for a 21st Century Society

1. Introduction

The founding fathers realized a census was necessary to establish the republican form of government based on the representation of the resident population as it distributed itself across the original 13 states and then spread into new territories, soon to become new states. The decennial census was a mechanism to apportion the House of Representatives based on each state's population size. It was early recognized that the massive effort to count everyone in the growing country and locate them at their permanent residence could provide additional information inexpensively, which gradually led to an enlarged census of the social, demographic, and housing characteristics of the total U.S. population.

The 2020 Census structure is a short household questionnaire that allows an internet response linked to the much more detailed sample-based American Community Survey (ACS) (see Box 1). Our project's premise is that although this model – and its predecessors- has and could continue to serve the nation, 21st-century conditions provide an opportunity for an order of magnitude improvement. The key conditions emphasized are data sources not previously available.

One such data source comes from the government itself -- administrative records (federal, state, and local). Of course, these records are not themselves new, and sector by sector (health, education, taxes, births & deaths, social security, crime, etc.), they have been constructively used in policies relevant to their respective sectors. Only recently, as described in more detail below, have we begun to systematically explore the analytic and interpretive power of cross-sector data linkage.

Then, of course, the enormous growth of digital data is primarily occurring in the private sector. A vast array of behaviors (social media) and products (the internet of things), and transactions (credit card swipe data) are routinely and constantly recorded in great detail. We are well aware of the many barriers to the effective use of this data source for the policy tasks historically assigned to the census and other survey-based data collection agencies of the government. We take note of the error rates tolerated by the commercial sector and the lack of temporal continuity. We recognize the challenges of privacy and confidentiality. These difficulties, and more, notwithstanding, there are already success stories – e.g., the 2020 census updated address file, not, as before, based on thousands of enumerators walking the street but on using high-quality digital imagery and related technological innovations.

The 21st-century model we have in mind involves multiple-sourced data products that, taken together, can build a new information platform for the nation. The decennial is a component and a key one. That is, on every April 1 that occurs in a year ending in zero, a 100% enumeration is taken, which places all people at their usual place of residence. It is this data product that reapportions the House of Representatives. Other basic traits -- sex, age, ethnicity, and race – may be included. That is a topic for discussion. This enumeration is designed so that all other sample surveys can use it as the standard.

Other government surveys remain, including the ACS, Current Population Survey (CPS), and dozens more. But now – depending – they would be integrated with administrative record data and with externally sourced data resulting in a data platform expected to provide demographic and geographic granularity and temporality not presently available and do so in a manner that protects privacy and confidentiality.

The Census Bureau has already made substantial progress in what we view as the primary task of the April 1-Year Zero census by building a comprehensive Statistical Administrative Records System. A description of this system appears in the November 3, 2016, Federal Register Notice (Department of Commerce 2016), which provides a Systems of Records Notices describing this valuable data resource.

To meet the country’s needs, we propose a new process whereby essential data will be built over the decade to provide a complete and accurate enumeration of the population of the United States for the first time in 2030. A vital component of this enumeration will be to meet the constitutional requirements of the 2030 Census to provide state population totals. This enumeration will also be used as a benchmark for tabulations created throughout the decade by the augmentation of the 2020 decennial census with multi-source federal and private sector sources. We call this an Essential-Data Count of the Resident Population coupled with a 21st Century Census Information Platform.

This concept paper describes the scientific innovations of decennial censuses and explains why this proposal is the next innovation. This is followed by a discussion of whether there are Constitutional or legal barriers to the Essential-Data Count of the Resident Population to meet the Constitutional requirements for apportionment. The details of how this would be done will be reported on as progress is made by technical working groups. Here we express the key task: design a 2030 census true to the broad purposes embedded in the Constitution that also continues a 240-year history of scientific and technical innovations that assure accuracy, privacy, coverage, detail, and timeliness within budget limits set by Congress.

2. Decennial Census Innovations

The Census Bureau improves the timeliness and reliability of the decennial census by incorporating scientific and technological advances. This has also led to innovation spillovers into the economy. Table 1 provides examples of these innovations over time. These include:

- implementation of the mail census enumeration system in urban areas;
- the development of the TIGER (Topologically Integrated Geographic Encoding and Referencing) database (the first nationwide digital map of roads, boundaries, and water), which led to the growth of the GIS industry; and
- implementation of the American Community Survey (ACS) that provides annual data to replace the decennial census long form that provided data once a decade.

For the 2020 census, the Census Bureau plans to increase the use of administrative records and third-party data, optimize self-response (including Internet response), and reengineer address canvassing and field operations to collect information from households that do not self-respond (US Census Bureau 2018b).

Even with this continuous flow of innovations, many challenges remain in conducting a census in the face of a rapidly changing and dynamic resident population. New conditions in the private and public sectors present an opportunity in 2030 to blend administrative, survey, and private sector data in the decennial collection and the continuous ACS. Science and technology innovations allow for the capture of such data through IT mechanisms that most individuals use today. These innovations have multiple reciprocal benefits across the public-private boundary; the decennial remains anchored to its public good responsibility and committed to accuracy, credibility, confidentiality, and independence immune to political or commercial interference. It retains its capacity to benchmark the representativeness and accuracy of other sources of national numbers, including those being constructed from administrative and third-party sources.

A clear benefit of the proposed Essential-Data Count of the Resident Population for the decennial 2030 is the containment of costs and minimizing the use of taxpayer dollars while collecting and publishing the highest quality data. This proposal for 2030 will smooth out budget cycles and optimize measurement per dollar spent.

Census data retains its responsibility for foundational benchmarking other data collection efforts, public and private. The proposed 2030 census will benefit from and build on numerous instances of blending survey, administrative, and private data that are underway. However, survey data will still be needed and will focus on variables that are not readily available from administrative data sources (e.g., hard-to-count populations).

An impressive history of innovation notwithstanding, additional improvements to accuracy may require a shift in the design of the Census. This proposal recommends changing from household-centric data collection to direct enumeration of individuals. There is active research on using other sources of federal, state, local, and private data. Government agencies are experimenting with obtaining company data through application programming interfaces (APIs), thus reducing the burden on the ACS and related surveys. The Census Bureau works with companies to implement methods and technology to perform statistical computations using company data (Jarmin, 2019).

Currently, the Census Bureau creates annual population estimates using aggregate administrative records at the county, state, and national levels, and the American Community Survey, starting with the census population estimates each decade. These administrative records include National Center for Health Statistics Vital Statistics to measure births and deaths; Internal Revenue Service, Medicare, and Social Security data to measure domestic migration, and the American Community Survey to measure migration to and from the United States (U.S. Census Bureau, 2019)

As the data user community becomes larger and more diverse, it will require the Census Bureau and other statistical agencies to engage with these users to learn about their measurement priorities and to create new products to meet their needs. Advocacy organizations will expect data relevant to their communities, for example, the urban homeless or population groups especially vulnerable to coastal flooding.

As the Federal Statistical System evolves to using a blend of administrative, survey, and private sector data, staying true to these core principles is even more critical. This topic will be addressed extensively as the methodological details for Census 2030 are developed and become codified in the 2030 operational plans. These principles are (NASEM, 2017):

Principle 1: A federal statistical agency must be in a position to provide objective, accurate, and timely information that is relevant to issues of public policy.

Principle 2: A federal statistical agency must have credibility with those who use its data and information.

Principle 3: A federal statistical agency must have the trust of those whose information it obtains.

Principle 4: A federal statistical agency must be independent from political and other undue external influences in developing, producing, and disseminating statistics.

3. Constitutional Requirements, Court Cases, and Legislative Review

This opportunity to continue innovations through a new approach to the decennial census is obtainable given the maturity in data science methods and technology and yet simultaneously revolutionary. The Census Bureau is already on a path to realizing this proposal given the use of the internet and administrative records in 2020 (JASON 2016, US Census Bureau. 2018b). Still, the shift being proposed requires answers to its constitutionality or other plausible legal challenges.

The U.S. Constitution requires a decennial enumeration of the population to apportion the House of Representatives seats to each state:

The actual Enumeration shall be made within three years after the first meeting of the Congress of the United States, and within every subsequent Term of Ten Years, in such Manner as they [Congress] shall by Law direct (U.S. Constitution (Article 1, Section 2).

Section 2. Representatives shall be apportioned among the several States according to their respective numbers, counting the whole number of persons in each State ... Amendment XIV (Ratified July 28, 1868)

The Constitution does not prescribe the process to be used to create the enumeration, except that an “actual” enumeration is to occur every ten years.

In addition to the Constitution, the U.S. Census Bureau, and hence the decennial census, is shaped by legislation.

The first is the Act to Provide for a Permanent Census Office (1902) that made the Census Office a permanent agency. The second is the Census Act - U.S.C. Title 13 (1954) that combined existing laws governing the Census Bureau's statistical programs and activities and codified them in Title 13. It does not specify which subjects or questions to include in the census nor who is to be counted. It does specify other provisions related to confidentiality and sampling described below.

Title 13, Section 6 authorizes the Census Bureau to use information from other Federal departments and agencies and to acquire reports from other governmental and private sources.

Title 13, Section 9 directs that information collected by Census Bureau be kept confidential. Address lists directly associated with the enumeration, even an address list without names, is considered “information as confidential.” Section 195 allows some uses of sampling, as stated: “Except for the determination of population for purposes of apportionment of Representatives in Congress among the several States, the Secretary shall, if he considers it feasible, authorize the use of the statistical method known as “sampling” in carrying out the provisions of this title.”

Supreme Court and lower court cases about the decennial census have generally focused on challenges about using sampling procedures proposed for adjusting the census for apportionment. There have also been challenges to the Census Bureau processes regarding whom to include in the decennial census, such as the military located overseas and illegal immigrants. (Table 3 presents selected court cases related to the decennial census and apportionment.) The Courts have generally upheld the Census Bureau’s authority to determine enumeration procedures, except sampling to adjust for the undercount for apportionment. Since the court explicitly prohibited only sampling for reapportionment, the interpretation is that sampling can be used for adjusting the counts for redistricting and other statistical purposes but not for apportionment.

From a Constitutional and legal perspective, barriers do not appear to exist to implement an *Essential-Data Count of the Resident Population* for the 2030 decennial census, coupled with a *21st Century Census Information Platform*.

4. Conclusions

This concept paper describes an *Essential-Data Count of the Resident Population* for the 2030 decennial census, coupled with a *21st Century Census Information Platform*. Broad consensus will be needed around this proposal to move forward involving bipartisan and advocacy groups, businesses, residents, and ultimately Congress.

This concept paper is the first step toward developing that consensus.

Box 1. American Community Survey

The American Community Survey (ACS) is an ongoing survey that annually provides information to examine changes about our society and its people. Started in 2005, it replaced the long form on the decennial census for 2010. The ACS provides annual data that is more accurate, timelier, and geographically granular than the traditional long-form decennial data that was administered to a sample of the population.

The ACS has an annual sample of approximately 3.5 million housing unit addresses and group quarters in the United States. The Census Bureau selects a random sample of addresses to be included in the ACS. Each address has about a 1-in-480 chance of being selected in a given month, and no address is selected more than once every five years. Data are collected by internet, mail, telephone interviews and in-person interviews. Approximately one-third of those who do not respond to the survey by mail or telephone are randomly selected for in-person interviews. About 95 percent of households across all response modes ultimately respond.

Information from the ACS generates data that help determine how federal and state funds are distributed each year. The ACS also provides data about jobs and occupations, educational attainment, veterans, whether people own or rent their homes, and other social, housing, and demographics topics. Public officials, planners, businesses, and researchers use this information to plan for hospitals and schools, support school lunch programs, improve emergency services, build bridges, and inform businesses looking to add jobs and expand to new markets, and more.

Source: Census (2018a)

Selected Background Materials

Anderson, M. J., (2015) *The American Census, A Social History*, 2nd edition, Yale University Press.

Anderson, M. J., Citro, C., & Salvo, J. J. (2011). *Encyclopedia of the US Census: from the Constitution to the American Community Survey*. CQ Press and Sage.

Anderson, M. J. & Feinberg, S. (1999). *Who Count? The Politics of Census – Taken in Contemporary America*. Russell Sage Foundation. New York.

Department of Commerce (2016). Notice of Amendment, Privacy Act System of Records; Commerce/Census-8, Statistical Administrative Records System (November 3, 2016). 81 FR 76554. Pages: 76554-76557. Federal Register: The Daily Journal of the United States.

Hillygus, D. S., Nie, N. H., Prewitt, K., & Pals, H. (2006). *The hard count: The political and social challenges of census mobilization*. Russell Sage Foundation.

Jarmin, R. S. (2019). Evolving measurement for an evolving economy: Thoughts on 21st Century US economic statistics. *Journal of Economic Perspectives*, 33(1), 165-84.

JASON. (2016). *Alternative Futures for the Conduct of the 2030 Census*. JSR-16-Task-009. November.

Kish, L., & Fellegi, I. P. (1981). *Using cumulated rolling samples to integrated census and survey operations of the Census Bureau: An analysis, review, and response*. US Government Printing Office.

National Academies of Sciences, Engineering, and Medicine. (2017). *Principles and Practices for a Federal Statistical Agency*, Sixth Edition. Washington, DC: The National Academies Press.

National Research Council. (2004). *The 2000 Census: Counting Under Adversity*. Washington, DC: The National Academies Press.

National Research Council. (2004). *The 2000 Census: Counting Under Adversity*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/10907>.

Prewitt, K. (2010). What is political interference in federal statistics? *The ANNALS of the American Academy of Political and Social Science*, 631(1), 225-238.

Prewitt, K. (2000). The US decennial census: Political questions, scientific answers. *Population and Development Review*, 26(1), 1-16.

US Census Bureau. (2016). Population and Housing Unit Estimates. Demographic Analysis Last Revised: December 8, 2016

US Census Bureau. (2018a). About the American Community Survey. Last Revised: June 17, 2018.

US Census Bureau (2018b). 2020 Census Operational Plan, A new Design for the 21st Century. Version 4.0.

US Census Bureau (2019). State Population Totals and Components of Change: 2010-2019

Appendix Tables

Table 1. Selected Decennial Census Innovations

1890	Pioneered machine tabulation
1940s	Established concept of measurement error in federal surveys introduced sample surveys to collect data between censuses
1950s	Used first nondefense computer; introduced sample surveys to collect data between censuses; initiated quality control methods , to measure accuracy of census including Post-Enumeration Survey to assess quality and procedures; research on Demographic Analysis to pinpoint where under/overcounts existed but could not identify which individuals were missed or why
1960	Systematically measured the undercount through Post-Enumeration Survey and Demographic Analysis
1970	Implementation of the short census form to collect basic information from 5/6 th of the population and the long census form to collect more detailed social, demographic, and housing data from a sample of 1/6 th of the population. Increased emphasis on accuracy based on 3 decades of earlier research; moved to mail census enumeration system for 60% of population, mostly in urban areas
1980	Development of automated geographic files , the first computerized map of the US, called TIGER ; no adjustments were made to data in all reports (and uses); undercount remained controversial; Kish Report introduces idea of continuous measurement that led to ACS implementation in 2005
1990	Shelter and Street Night (S-Night) was established as a concentrated effort to enumerate persons experiencing homelessness . It resulted in the Census Bureau partnering with organizations and localities to develop a supplemental database prior to conducting the Census .
2000	Census Bureau implemented Accuracy Coverage Evaluation (ACE) program to measure and potentially statistically correct the census for undercounts and overcounts. In addition, using enhanced marketing of the census , mail responses were improved relative to expectations, and undercounts of traditionally hard-to-count populations were reduced.
2005	Introduction of American Community Survey (ACS) to replace decennial census long form
2010	Streamlined census to “10 questions in 10 minutes” (promotional campaign to complete the census); implemented Census Coverage Measurement program to estimate coverage problems. Implemented an extensive study of how administrative records could replicate the census counts.
2020	Reengineering field operations, making fuller use of government administrative records and private data sources , optimizing self-response (including Internet response), and reengineering address canvassing and the field operations to collect information from those households that do not self-respond

Sources: Anderson, Citro, & Salvo (2011); Anderson (2015); U.S Census Bureau (2018b)

Table 2. Constitutional and Legal Requirements Related to Use of Decennial for Apportionment and Related Activities

U.S. Constitution (Article 1, Section 2)	The actual Enumeration shall be made within three Years after the first Meeting of the Congress of the United States, and within every subsequent Term of ten Years, in such Manner as they [Congress] shall by Law direct ...
U.S. Constitution, Article I, section 2, clause 3	Section 2. Representatives shall be apportioned among the several States according to their respective numbers, counting the whole number of persons in each State ...
Amendment XIV (Ratified July 28, 1868)	The amendment grants citizenship to "all persons born or naturalized in the United States" which included former slaves who had just been freed after the Civil War.
Act to Provide a Permanent Census Office (1902)	Made the Census Office a permanent agency
Census Act - U.S.C. Title 13 (1954)	Combined the existing laws governing the Census Bureau's statistical programs and activities, codifying them in Title 13., and designates the Secretary of Commerce as overseeing the Census Bureau functions.
Census Act - U.S.C. Title 13, Section 6	Authorizes the Census Bureau to use information from other Federal departments and agencies and acquire reports from other governmental and private sources
Census Act - U.S.C. Title 13, Section 9	Directs that information collected by Census Bureau be kept confidential. Address lists directly associated with the enumeration, even an address list without names, is considered "information as confidential."
Census Act - U.S.C. Title 13, Section 195	Allows some uses of sampling, as stated: "Except for the determination of population for purposes of apportionment of Representatives in Congress among the several States, the Secretary (of Commerce) shall, if he considers it feasible, authorize the use of the statistical method known as "sampling" in carrying out the provisions of this title."
H.R. 4174 (115th): Foundations for Evidence-Based Policymaking Act of 2017	Each agency shall: (1) develop and maintain a comprehensive data inventory for all data assets created by or collected by the agency, and (2) designate a Chief Data Officer who shall be responsible for lifecycle data management and other specified functions. (3) recommends sharing of information between agencies as the default unless existing laws prohibit the practice.

Table 3. Selected Court Cases Related to the Decennial Census and Apportionment

Court Cases	Outcome
U. S. Department of Commerce v. Montana 504 U. S. 442, 1992, No. 91-860	Montana’s 900,000 residents in 1990 got only one representative in Congress while the average district elsewhere included about 600,000. This was not a constitutional violation.
Young v. Klutznick, 652 F. 2nd, 617 (Detroit, 1980) and Carey v. Klutznick, 508 F. Supp. 416 (NY city and NY state, 1980) – 2 of the 50 cases	50 lawsuits filed by local and state governments against the Census Bureau in 1980 to correct for the differential undercount. The courts initially sided with the plaintiffs, the decisions were later stayed, and the unadjusted 1980 results were released to the president on time.
Barbara Franklin, Secretary of Commerce, et al., Appellants V. Massachusetts et al. (1992), 505 U.S. 788. No. 91-1502.	Massachusetts challenged the inclusion of overseas military and federal government population in population counts. Court upheld the Census Bureau’s authority to determine counting procedures.
Dept. of Commerce v. House of Representatives (1999), 525 U. S. 316, 1999, No. 98-408.	Sampling and statistical methods may not be used to adjust the apportionment population for net census undercount.
Utah, et al., Appellants v. Donald L. Evans, Secretary of Commerce, et al. (2002), 536 U.S. 452 (2002) 182 F. Supp. 2d 1165, affirmed. No. 01-714	Utah challenged the Census Bureau’s practice of counting overseas population and its use of imputation procedures. Utah claimed that the use of imputation in the 2000 Census was sampling and therefore could not be used for apportionment, The Court (1) upheld the Census Bureau’s authority to determine counting procedures; (2) determined that for enumeration of the population, the use of imputation was not sampling and therefore is allowed. (See note below).

Note: After repeated attempts to contact a household that did not return its census form, the Census Bureau can use imputation to infer that the address or unit about which it is uncertain has the same population characteristics as those of its geographically closest neighbor of the same type. Imputation increased the total population count by about 0.4% but the distribution was not equal across the states, e.g., 0.4% in North Carolina and 0.2% in Utah in 2000.