

## Protecting the Accuracy of the



Public distrust of government, vulnerable computer systems, a possible citizenship question, and new privacy protection techniques are formidable challenges confronting the decennial count.

Census Day, April 1, 2020, is when everyone in the United States will be asked to answer questions on the 24th decennial enumeration of the nation's population, the latest in an unbroken, constitutionally mandated tradition that began in 1790. Like its predecessors, the 2020 census will be the nation's largest peacetime mobilization, employing an army of enumerators to follow up with households that do not respond either online (an innovation in 2020) or by mail, in order to obtain as complete a count as possible. It will also be conducted under tight time constraints, with statutory mandates to provide state population counts to the president by December 31, 2020, for use in reapportioning seats in the US House of Representatives, plus small-area (block-level) population data by age and race/ethnicity to the states by March 31, 2021, to use in redistricting congressional seats. In addition, census data have a myriad of uses that are built into the fabric of the nation's society, economy, and polity:

- federal agency allocation of billions of dollars of federal funds to states and localities;
- business decisions about locating workplaces and projecting product demand;
- academic research on trends in family structure and living arrangements, migration, and the racial/ethnic composition of the population;
- state and local government location of neighborhood services and schools; and
- support to key social and economic statistics such as the unemployment rate, poverty rate, and consumer price index.

The Census Bureau has projected that the 2020 census will cost \$15.6 billion over the 10-year cycle that began with planning in 2013 and will conclude in 2023 with the release of all data products and evaluation studies, including estimates of the accuracy of the count. The Bureau projects that the nonresponse follow-up operation will require hiring almost 500,000 part-time enumerators and other field operations staff, who will need to complete their work within a short time span.

Censuses in the modern era, beginning in 1970 when the Bureau turned to encouraging household self-response in place of in-person interviewing in census-taking, have delivered data products for reapportionment, redistricting, and other uses on time, if not always within the estimated budget. They have also delivered acceptable levels of accuracy, with some censuses being better than others. But experts have raised serious concerns about the 2020 census, and some have warned of an impending disaster. A number of factors are contributing to this sense of alarm:

***Distrust of the federal government and the push to include citizenship on the census.*** The Trump administration's effort to include a question on citizenship, in combination with heightened distrust of the federal government, has led to fears that many immigrants, whether or not they are citizens, will not respond to the census. The Census Bureau, of course, will follow up with nonresponding households, but it may have to estimate the number of people in many more households than is typically the case. And even with its best efforts, there may be a larger net undercount of the total population. Further, there may be a larger differential undercount between minorities—particularly people of Hispanic origin, many of whom are immigrants—and others, with consequences for

political representation and allocation of governmental resources. Three US district courts have ruled that the secretary of commerce violated the Administrative Procedures Act and therefore should not be able to add a citizenship question, and two of those courts further concluded that adding the question is a constitutional violation. The US Supreme Court heard oral arguments on the matter and is expected to render a decision by the end of June 2019, barely in time for the Bureau to finalize the questionnaires.

**Computing systems vulnerabilities.** The 2020 census will, for the first time, offer an online response option to a large percentage of the population, a practice already in use in many other countries. The Census Bureau routinely uses online response for its largest survey—the American Community Survey, which samples 250,000 households each month. Although online response affords many advantages, the challenge for the 2020 census is that it will be the first time that the Bureau will grapple with the sheer volume of near-simultaneous response encouraged by the wave of publicity surrounding the decennial count. There is a real possibility that the Bureau’s computers will experience a crash similar to the one that occurred in fall 2014 during enrollment under the Affordable Care Act. Another vulnerability is that a foreign government, or an individual or group of individuals, could seek to break into the Bureau’s computing systems and release confidential information. Such a breach could have devastating consequences for 2020 and future censuses and surveys by undermining trust that the Bureau can fulfill its legal requirement to protect the confidentiality of individual responses.

**Funding shortfalls that impeded innovation.** The 2020 census will incorporate significant operational and methodological innovations in at least four areas: use of the internet for self-response; use of smartphones by enumerators, not only to record answers from households but also to receive their daily workload of addresses to visit (accompanied by a significant reduction in local offices and staffing); use of online tools to update the Master Address File; and use of administrative records to reduce the number of visits by enumerators to some kinds of households. These are major steps forward toward a more cost-effective census, but budget shortfalls in 2017 and 2018 and the federal government shutdown in early 2019 impaired the ability of the Census Bureau to thoroughly test and fully exploit these innovative strategies.

**Introduction of new privacy protection techniques for important data products.** In response to the very real threats to the confidentiality of census information from the quantum advances in computing technology and the availability of data about people on the internet, the Census Bureau is planning to employ new data protection methods

for its products. These cutting-edge “differential privacy” techniques, developed in the past 15 years by computer scientists, introduce controlled noise into the estimates. Data users, however, are concerned that in navigating the trade-off between data protection and data accuracy, the scales may tilt too much against accuracy.

This is not the first census to face significant challenges, and a review of how these difficulties were addressed could offer lessons on how to manage the 2020 census.

## A brief history

The US decennial census of population traces its roots to the colonial period. The British government, eager to know the extent of population growth and consequent economic development that could benefit the mother country, initiated a large number of population censuses in its American colonies from the early 1600s to the revolution. Data collected often included age, sex, and race—the latter of interest to distinguish those eligible for military service and taxes—and sometimes family and household structure. In all, 46 censuses were conducted in nine of the original 13 colonies.

The functions and operations of censuses were therefore familiar to the delegates at the 1787 Constitutional Convention. The US Constitution became the first document to mandate a population census and give it a fundamental role—namely, as the means to regularly and peacefully reallocate power and resources as the nation grew and people migrated to different parts of the country. Article I, Section 2, of the Constitution reads, in part: Representatives and direct Taxes shall be apportioned among the several States which may be included within this Union, according to their respective Numbers, which shall be determined by adding to the whole Number of free Persons, including those bound to Service for a Term of Years, and excluding Indians not taxed, three fifths of all other Persons. 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 shall by Law direct.

Note the infamous injunction to count “three-fifths of all other Persons”—that is, slaves. This wording represented a compromise. Northern states did not want slaves counted for purposes of reapportionment of seats in the House but were willing to have them counted for direct taxation of the states by the federal government, and vice versa for southern states. James Madison thought the three-fifths compromise also would protect the integrity of the census, as he explained in *The Federalist No. 54*: “The States should feel as little bias as possible to swell or to reduce the amount of their numbers.... By extending the rule to both [taxation and representation], the States will have opposite interests which will control and balance each other and produce the

requisite impartiality.”

The 14th Amendment to the Constitution, passed in 1868, made all persons born in the United States (including former slaves) citizens of the United States. It also specified that reapportionment of the House would be based on “the whole number of persons in each State, excluding Indians not taxed,” rendering the three-fifths clause moot.

Note the continued exclusion of “Indians not taxed,” a group that was never well defined but essentially taken to be Indians living on reservations or, in the instructions to enumerators in the 1880-1910 censuses, “roaming individually, or in bands, over unsettled tracts of country.” A 1924 statute mooted the need for a definition by providing that “all noncitizen Indians born within the territorial limits of the United States be, and they are hereby, declared to be citizens of the United States.”

The decennial census has been characterized by both change and continuity. Among the key changes, marshals on horseback have given way to temporary enumerators working for a permanent Census Bureau. Enumeration methods based solely on personal visits have given way to primary reliance on self-response. Data capture and processing technology has changed from handwritten ledgers to Hollerith punch cards to the first UNIVAC computer to modern integrated computing systems. Questionnaire content has changed from a handful of questions in 1790, to 32 questions in 1910, to asking some questions of a sample of the population beginning in 1940 (becoming the “long-form” questionnaire in 1970), to moving the long-form content to the continuous American Community Survey in 2005. The 2000 census asked only a handful of basic questions on date of birth, race, ethnicity, sex, household relationship, place of residence, and whether the home is owned or rented. Evaluation of the accuracy of the census has evolved from anecdote (George Washington and Thomas Jefferson opined that the “real numbers” in 1790 greatly exceeded the official counts) to state-of-the-art statistical methods for estimating net undercount and differences in net undercount rates among population groups and geographic areas. (See the text box.)

Other aspects of the census have persisted. There is the same standard for determining where residents should be counted in 2020, namely, “usual residence,” as the “usual place of abode” standard specified by the first US Congress for the 1790 census. With a few exceptions, such as US citizens living abroad, who have been counted in some censuses and not others, the broad mandate remains to count all residents, regardless of age, sex, race, citizenship status, or other characteristic—excepting only citizens of other countries who are part of a diplomatic compound or temporary visitors to the United States. The

## What Are Net and Differential Census Undercount Rates and How Are They Estimated?

The Census Bureau knows that it is impossible to count every individual in the country, so in recent censuses it has used what is called dual-system estimation to construct estimates of the total population and net undercount rates, expressed as 100% minus the ratio of the census count to the dual-system estimate. For example, if the ratio were 98%, then the net undercount rate would be 2%. Differential undercount is the difference between the net undercount rate for two population groups (say, non-Hispanic whites and African Americans).

Dual-system estimation was developed from so-called capture-recapture methods that are used to estimate the size of animal populations. For example, to estimate the number of fish in a lake, a sample of fish are captured, then tagged and returned to the lake; later, a second sample of fish are captured, some of which will have tags and others not. It is assumed that the ratio of tagged fish to total fish in the second sample equals the ratio of the first sample to the total in the lake. If 200 fish were captured in the first sample, and tagged fish comprised 25% of the second sample, then it can be calculated that there is a total of 800 fish in the lake.

For a census of population, it is not possible to directly “tag” individuals. Also, there are erroneous enumerations in the census (e.g., people born after Census Day) and people who were imputed into the count because no respondent could be located at a valid address. Dual-system estimation uses a post-enumeration survey as the second “system.” In this survey, an independently drawn sample of blocks and their addresses are visited after completion of census nonresponse follow-up, and respondents are asked where they lived on Census Day. The results are then matched to census records in the same blocks to determine both the number of matches and the number of correct census enumerations to use in the dual-system formula.

legislative act that authorized the 1790 census also made response to census inquiries mandatory, as is still the case today.

### The modern census

Modern census-taking began in 1970, which was the first census to use self-response as the primary enumeration method, by developing an address list (later dubbed the Master Address File), mailing out questionnaires, and asking respondents to complete them and mail them back. Most households received the short-form questionnaire; 15% of households received a somewhat more detailed questionnaire; and 5% received the longest form, which asked people born abroad whether they were naturalized citizens, aliens, or born abroad of American parents.

The 1970 census was also the first census to distribute data products in computer form (the Census Bureau pioneered computer processing of questionnaire responses in 1950 and 1960) and the first census conducted following passage of the 1965 Voting Rights Act, which was a key driver for census data at the block level for the entire nation to use in redistricting to meet court standards for compact, contiguous, and equal-population districts. The legislation also requires the Bureau to designate jurisdictions (counties and county subdivisions) that must make voting accessible to voting-age citizens not proficient in English, which the Bureau does by using data from the American Community Survey (previously from the long-form sample).

Below I briefly list key problems for each modern census, from 1970 through 2010; how each problem was resolved; and the bottom line of that census in terms of timely delivery of data products, costs, and self-response rates (see Figure 1), and estimated total and differential net undercount of the population (see Figure 2).

**Figure 1 Short-Form and Long-Form Mail Return Rates and Costs per Housing Unit (2018 Dollars), 1970–2010 Censuses**

		1970	1980	1990	2000	2010
Mail Return Rates	Short Form	87%	81%	75%	78%	79%
	Long Form	86%	80%	71%	71%	NA
Costs per Housing Unit		\$20.40	\$39.11	\$50.87	\$83.16	\$118.94

**Figure 2 Net Population Undercount, 1970–2010 Censuses**

	1970	1980	1990	2000	2010
Total	2.71%	1.22%	1.65%	0.12%	-0.01%
Black	6.51%	4.50%	5.52%	2.78%	NA
White & Other	2.21%	0.77%	1.08%	-0.29%	NA

Note: Net undercount rates measured with demographic analysis (DA); preferred DA results not available for 2010 (PES methodology not available for 1970 or 1980). Minus sign indicates a net overcount.

### 1970

*Last-minute demand for a new question.* An interagency group under the Bureau of the Budget pressed for a question on Hispanic origin in early 1969. (The current legal requirement that census topics and questions be finalized and shared with Congress three years and two years prior to Census Day, respectively, became law only in 1976.) The solution was to add a question to the 5%-sample long form but not the 15%-sample long form, thereby minimizing costs to revise questionnaires. The question itself was poorly designed. Many residents in the Midwest, for example, said they were “Central or South American.” Subsequent censuses asked everyone about Hispanic ethnicity.

*Too few households.* Real-time feedback from the nonresponse follow-up operation revealed too many housing units being classified as vacant or nonresidential. The solution was to revisit a sample of these units and use the results to probabilistically reclassify some units as occupied households. Ultimately, sampling for this purpose was forbidden, first by the Census Bureau in the 1980 and subsequent censuses, and then by the Supreme Court in 1999 for the state population counts used for reapportionment.

*Upshot.* The 1970 census cost less per housing unit than subsequent censuses and had a high mail-response rate. It delivered computerized products for mandated and other uses. It had a higher net undercount rate and a higher differential undercount rate between African Americans and others than subsequent censuses, although its rates were lower than in earlier censuses.

### 1980

*Concern about differential undercount of minority groups in the context of the civil rights movement.* Demand grew in the 1960s and 1970s for improved coverage of minority groups in the census and for the Census Bureau to use statistical methods to adjust the census results to minimize coverage errors. The Bureau put into place coverage-boosting programs, such

as a complete recheck of vacant housing units, a “Were You Counted” campaign, and a cross-check with driver’s license lists and other records, and conducted a post-enumeration survey (PES) to estimate census coverage using dual-system estimation. However, the Bureau ruled out statistical adjustment of the 1980 census unless ordered to do so by courts.

*Flood at main computer center.* Two UNIVAC mainframes were destroyed and two others were damaged in August 1979 when the fire sprinkler system accidentally discharged in the main computer room of the Census Bureau’s headquarters, in Suitland, Maryland. The Bureau was able to minimize disruption to the 1980 census through temporary off-siting and rapid replacement of equipment, but the incident briefly fueled concerns about meeting census deadlines.

*Budget cut in 1981.* Data collection was not affected, but data processing had to be stretched out. The solution was to delay products from the long-form sample.

*Upshot.* The 1980 census cost more per housing unit in real terms than 1970. Spending on coverage improvement had a payoff in reduced net and differential undercount. Nonetheless, states and localities filed lawsuits demanding statistical adjustment, which dragged on for several years before being denied.

## 1990

*Arguments over content.* The Office of Management and Budget wanted to reduce the content on both short and long forms. Congress wanted more detailed categories for Asian Americans in the race question on the short form. After an outcry from data users, most content items were retained, and more Asian-African categories were added to the race question.

*Unexpectedly steep drop in mail response rates.* Response rates to censuses and surveys began declining in the United States and other countries in the 1970s. The Census Bureau planned for a decline in 1990 from 1980, but not for the extent of decline that occurred (see Figure 1). For example, it had no plans to send a second questionnaire to nonresponding households. The solution was to allot more funding and more enumerators to nonresponse follow-up.

*Lawsuits seeking statistical adjustment of census results for net undercount.* The Commerce Department preemptively announced in October 1987 that it did not intend to statistically adjust the 1990 census, which touched off a number of lawsuits brought by states and localities as early as 1988, demanding adjustment. The principal suit was stayed, and an injunction against conducting the 1990 census at all was lifted after Commerce agreed to reconsider adjustment. Ultimately, the 1990 census proved to have a larger overall and differential net undercount than 1980. The Census Bureau

advised in favor of statistical adjustment, but the then secretary of commerce, Robert Mosbacher, declined to do so in July 1991, and the courts upheld his decision.

*Upshot.* The Bureau director for 1990 characterized that census as a “technological triumph and public relations disaster.” The technology for mapping and data collection worked, but the cost per housing unit was higher than in 1980, primarily due to lower mail response, and coverage was estimated to have been worse than in previous censuses. Planning to “reengineer” the census began earlier than usual for 2000, and there was a push for pilot work on what became the American Community Survey to lift the burden of the long-form sample from the census.

## 2000

*Competing designs in play.* The Census Bureau decided in the mid-1990s to pursue a design that would use sampling for nonresponse follow-up to reduce costs and improve accuracy, and to conduct a post-enumeration survey that would be used for statistical adjustment of the census results in time to deliver the required data products for reapportionment and redistricting. The congressional Republicans who took power in 1994 wanted a traditional census. The matter was appealed to the courts, so the Bureau could not finalize the design or build out the necessary computing systems until the Supreme Court ruled. It finally did so in January 1999, finding that the sampling-heavy design violated census law against using sampling to generate apportionment counts. The court did not rule on whether sampling was permissible for other uses of census data, such as redistricting, and deliberately declined to opine on whether the sampling-heavy design was constitutional. The Bureau went into overdrive on software development, putting systems into operational use with minimal testing.

*Lower-than-expected response for respondents sent the long form.* Partly due to politicians’ unguarded comments about the perceived intrusiveness of long-form questions, mail response from those receiving the long form was lower than expected, and the Census Bureau again had to devote more money and enumerators to nonresponse follow-up. Short-form response did not decline from 1990, due to the extensive use of paid advertising (instead of public service announcements) and promotional partnerships with many organizations and localities.

*Optical-character reader (OCR) software not up to the workload.* For the first time in a census, the Census Bureau issued a contract for a vendor to process mailed-in questionnaires through an OCR system rather than using machinery originally built in the basement at Bureau headquarters. In testing, the software was not able to handle the volume of responses in a timely manner. The solution was to process just the short-form information first and to delay processing the long-form information.

*Testing that did not take feasibility into account.* To boost self-response, the Census Bureau tested a second questionnaire mailing with promising results. But investigation of whether vendors could process the second questionnaires within the time and at the scale required did not occur until late in the decade. A second mailing proved infeasible, so that idea was scrapped.

*Local population estimates suggested a high rate of duplicate enumerations.* At the height of data processing, it appeared that duplicates in the Master Address File were resulting in duplicate enumerations. Software was written to match all responses and drop likely duplicates from the count.

*The post-enumeration survey overestimated the undercount.* Comparison with another coverage evaluation method (demographic analysis) revealed that the survey suffered from the same problem of duplicate enumerations as the census. The solution was to reinterview survey respondents and carefully reestimate each component of the undercount analysis. The Census Bureau and outside experts, based on this experience, concluded that statistical adjustment of census data for redistricting and other uses could not be carried out in a timely manner or with sufficient accuracy. It was helpful for acceptance of this decision that the 2000 census achieved improved coverage.

*Upshot.* The 2000 census saw a continued increase in per housing unit cost, no further decline in response for short forms, poor quality of long-form data in terms of high missing responses for many individual items (another motivating factor for the American Community Survey to take over the long-form content), and improved total and differential coverage.

## 2010

*Failure to adequately specify and oversee a large contract for census operations.* With the long form moved to the American Community Survey and undercount adjustment off the table, all signals pointed to a smooth 2010 census operation. The Census Bureau issued a contract to supply handheld devices (predecessors to smartphones) to enumerators to use in address canvassing and nonresponse follow-up, together with the necessary operational control software to link headquarters, regional and local census offices, and enumerators. The Bureau, however, did not give the vendor enough information about all the complexities of census-taking (e.g., that housing ranges from single-family unattached homes to apartment complexes with hundreds of units) and did not issue the contract early enough to allow for testing of the vendor's hardware and software. It became clear at the end of 2008 that that hardware and software were not up to the job. The solution was to limit the contractor's scope of work, scramble within the Bureau to develop an operational control system, conduct

nonresponse follow-up with paper questionnaires, and pour a lot more money into all operations.

*Upshot.* The 2010 census cost even more per housing unit than any of its predecessors, but achieved a good mail response rate, and a good count—low net undercount overall and low differentials between minorities and others. However, the chaotic last-minute planning necessitated by the meltdown of the contractor's systems meant that 2020 planning was not able to build on 2010.

## Lessons learned

A review of the past five censuses highlights some persistent problems. Most of these censuses experienced difficulties with the quality of the Master Address File and the process of constructing it, including the pre-census operation in which enumerators walked every block in the country to check and add to the address file. Most experienced information technology problems of some sort. All experienced the decline in response rates that has affected censuses and surveys worldwide. All experienced concerns about differential undercount of minorities, and most had lawsuits calling for statistical adjustment for net and differential undercount or for other "corrections." For example, Utah sued in 2000 to have Mormon missionaries abroad counted in the census, but it lost the case. The 2000 and 2010 censuses experienced problems with contract management and the quality of contractor-provider hardware and software.

Three actions helped these censuses resolve these persistent problems:

- Deployment of statistical expertise to improve coverage evaluation methods, match enumerations to detect duplicates, and estimate misclassifications of vacant units from a sample.
- Reversion to traditional methods, such as the paper-based nonresponse follow-up used in 2010, and in-house software systems.
- Expansion of programs to improve coverage and use of paid advertising and partnership programs to boost self-response.

In retrospect, the past five censuses were a mixed success. All met statutory deadlines for key data products. All achieved reasonably good coverage of the population, with particular success in 2000 and 2010. The price of better coverage and the need to overcome potentially existential challenges were sharp increases in costs per housing unit over the period. The challenges faced by each census impaired the learning curve for the next. Census planning would start fresh each time, but fall back on old methods when new ideas did not pan out. There was not enough cumulative learning for evidence-based planning decisions.

## Protecting 2020

The fact that every modern census has delivered acceptable results on time is reassuring for 2020, but it should not be forgotten that the dedication and extraordinary hard work of Census Bureau staff have had to save the day repeatedly. Yet the two biggest challenges for 2020—a poisoned political environment, which can only be exacerbated by the addition of a citizenship question, and the threats of cyberattacks during enumeration and processing—give me and others cause for concern. Here are four suggestions for additional steps the Census Bureau could take to protect 2020:

**Hardening information technology systems.** Should there be cyberattacks or other serious problems with 2020, it will not likely be good enough to deploy in-house information technology personnel from other parts of the Census Bureau, excellent as these folks are. What could make sense is to take preemptive action by bringing in SWAT teams from internet and computer giants now. One team could focus on internet security; a second team could focus on a deep dive into all the processing systems to be sure they are interoperable at scale. Also, any temptation to add bells and whistles to operating systems that are not absolutely necessary should be resisted.

**Addressing the citizenship question.** Many members of the general public support asking a citizenship question in 2020, whereas many others are frightened of the implications for a complete count of immigrant communities and minorities generally. Citizenship has never been used for legislative redistricting, although there are state legislatures that have indicated an interest in doing so. Citizenship data have never been published at the block level. Block statistics did not exist before the 1940 census, the citizenship information collected from everyone in 1940 and 1950 was not published for blocks, and beginning in 1960 collection of citizenship information was limited to the long form and later the American Community Survey.

Should the question be added, I believe the Census Bureau should decline to release citizenship data at the block level. The data could be protected by the new differential privacy techniques, but the utility of the block-level data for redistricting and other purposes could be significantly impaired. The reason is the amount of noise that would likely need to be injected into the data to protect against reidentification of individuals, given the addition of a citizenship variable to the other census characteristics. Whether or not the question is added, the Bureau should decline to construct data on citizenship status by matching Social Security and immigration records with (or without) 2020 census information. The Bureau has the technical capability and

is mandated by law to protect such matched records and not release them for law enforcement purposes. Nonetheless, if it were to engage in such matching, the Bureau could take a blow to its reputation as a purveyor of objective information for the common good and be perceived as violating a fundamental principle for a federal statistical agency. Any such matching operation should be the responsibility of the Social Security Administration or the Department of Homeland Security and should not include census data.

**Retaining promotional partners.** The Census Bureau deserves praise for its efforts to double down on its advertising, outreach, and partnership programs for 2020, which will be essential to overcoming fears in minority communities about responding. I believe, however, that many of the Bureau's current, and prospective, partners will be challenged in how wholehearted their support for the census can be should the citizenship question be included. Some organizations are already talking about asking people to respond to the census but skip the citizenship question. Indeed, all Americans will need to ask themselves what response is most appropriate to make.

**Preparing for incomplete response and lawsuits.** The Census Bureau must work with the scientific community to prepare for the likelihood of an impaired count and a spate of lawsuits. The Supreme Court has ruled out statistical adjustment for reapportionment but not for redistricting or other uses. Further, although the 2000 experience showed that a quality adjustment was not feasible within the statutory time frame for delivering block counts for redistricting, I expect that at least some states will opt to delay redistricting or redistrict again if they can use adjusted counts that more accurately represent their population.

This means that the Bureau must be fully prepared and resourced for two related operations: implementing procedures for imputing occupants to housing units that do not respond (in past censuses, this procedure accounted for 1.5% or less of the population) and if a citizenship question is added, for imputing missing responses to that question; and implementing procedures for evaluating the completeness of the count and for producing statistically adjusted counts should that be required. The Bureau should call on the resources of the scientific community to assist in vetting its methods and providing the transparency that will be essential for the Bureau and the 2020 census to be credible to policy-makers and the public.

*Constance F. Citro is a senior scholar with the Committee on National Statistics of the National Academies of Sciences, Engineering, and Medicine.*