

# THE NATIONAL ACADEMIES

*Advisers to the Nation on Science, Engineering, and Medicine*

Panel to Review the 2010 Census  
Committee on National Statistics  
Division of Behavioral and Social Sciences and Education

The Keck Center  
500 Fifth Street, NW  
Washington, DC 20001  
Phone: 202 334 3096  
Fax: 202 334 3751  
[www.national-academies.org/cnstat](http://www.national-academies.org/cnstat)

## *Prepared Statement of*

***Jack Baker***

***Senior Research Scientist, Geospatial and Population Studies,  
University of New Mexico***

***and***

***Member of the National Research Council Panel to Review the 2010 Census***

before the

Subcommittee on Federal Financial Management, Government Information,  
Federal Services, and International Security  
Committee on Homeland Security and Governmental Affairs  
U.S. Senate

July 18, 2012

Good afternoon, Mr. Chairman, Ranking Member Brown, and members of the subcommittee. Let me first thank you for the opportunity to testify before you today. I am Jack Baker, a member of the National Research Council<sup>1</sup> Panel to Review the 2010 Census. This panel was sponsored by the Census Bureau to provide feedback on methodological and operational aspects of the 2010 Census as well as to provide expert advice on how to design a more cost-effective 2020 effort. It is comprised of numerous experts in the fields of operations research, information technology, systems engineering, statistics, and demography—chosen to provide a broad perspective ranging among government, academia, and the private sector. In April 2011, the panel's chair (Dr. Thomas Cook) addressed this subcommittee and reviewed the findings of the panel's first interim report: *Change and the 2020 Census, Not Whether But How*. Today, I will update this subcommittee on developments since then that speak more directly to the Bureau's current planning efforts with respect to preparing for a more cost-effective 2020 Census. Some of these appear crucial, and worthy of strong Congressional support, for achieving a goal of successfully maintaining and even reducing Census costs without jeopardizing its quality.

I come to you from the perspective of an experienced demographic methodologist who has been involved in a number of pre-2010 data preparation programs as well as subsequent efforts to both evaluate the coverage of the census and to plan for improved future usage of

---

<sup>1</sup> The National Research Council is the operating arm of the National Academy of Sciences, National Academy of Engineering, and the Institute of Medicine of the National Academies, chartered by Congress in 1863 to advise the government on matters of science and technology.

spatial data resources (such as the Geographic Support Systems Initiative). As was the case when Dr. Cook testified to the subcommittee, I am sure that you understand that my expressed opinions in this testimony (and particularly in addressing any questions you may have) are my own, and should not be construed as formal guidance from either the panel or the Academies.

The panel's work on both reviewing 2010 and advising on 2020 planning is ongoing, but it is fair to say that the panel supports the basic conclusion from its first interim report: with eight years remaining before the 2020 Census, it is very possible for the Census Bureau to conduct the next census in a way that achieves large-scale reductions in cost (per housing unit) while maintaining quality. Last year, Dr. Cook testified about the necessity of attitudinal and perspective shifts on the Census Bureau's part in relation to this planning effort, and the need of a sufficient commitment to and prioritization of the planning process. Though I will not repeat it here, that guidance remains extremely valid.

In the panel's first interim report, we identified four priority topic areas for research and development, leading to an improved 2020 census:

1. The application of modern operations engineering to census field data collection operations;
2. Emphasizing multiple modes of response to the census, including response via the Internet;
3. Using administrative records-based information to supplement a variety of operations, and;
4. The continuous improvement and updating of the Bureau's geographic resources.

I would like to offer some general comments on the 2020 research and planning processes, and then focus on the last-mentioned of these topic areas.

From my perspective, the Census Bureau has been surprisingly receptive to the notion that ongoing testing, experimentation, and reformulation are important aspects of the 2020 planning process. The Bureau is moving to implement a more "adaptive" process of operational planning and field management for its data collection programs, to be tested and implemented first in the Bureau's major demographic surveys and then eventually to form the basis for 2020 census systems. I think that the Bureau's steps in this regard—developing a system in which survey respondents may be transferred to different response modes and approaches by interviewers, based on past contact attempts and contextual information—will allow the sort of exploratory thinking that our panel feels is crucial to addressing the challenges that lay ahead. I hope, and expect, that this same kind of approach may migrate into other census operations such as updating geographic resources like the Master Address File (MAF) or structuring field contacts in general (such as in nonresponse follow-up).

I think that this process of reformulating census-taking as a more organic, adaptive process rather than a string of only-loosely-integrated operations is a crucial step. For decades, the Bureau has tended to layer on more and more operations—often in the name of improving overall quality—without stepping back to consider costs, benefits, and cost-quality tradeoffs. (An earlier National Research Council panel made the same argument, with more supporting detail, in its 2010 final report *Envisioning the 2020 Census*.) I think that a management framework built on "adaptive design" can allow both a much more nimble consideration of the relationships

between census cost and census quality as well as a more detailed understanding of redundancies of effort (and resulting need to prioritize). Making this planning effort—and the research and development activities that will directly inform it—a funded priority with appropriately-committed Bureau resources is a must. In the long run, cost-savings and quality maintenance will only come with prior planning, and I believe that short-term up-front costs associated with this 2020 planning process is a worthwhile investment with major long-term cost offsets.

Effective means for updating the Bureau’s geographic data resources are a key aspect of any census design effort. I think it is very clear that shifts in response modes to the census (converting many, if not most, responses on paper questionnaires to Internet responses), coupled with a more adequate use of administrative records databases could reduce census costs considerably. But I think it equally clear that those gains would be undermined if the Bureau’s geographic resources were not up to par. Census respondents must be linked to specific physical locations (for purposes of apportionment and redistricting), and this requires high-quality spatial data resources that allow individual census forms to be linked to precise geographic coordinates.

One major element in 2020 census planning is a choice that the Census Bureau will have to make in the middle of this decade, and that is the extent to which it will conduct Address Canvassing prior to the 2020 census. As you know, the Bureau conducted Address Canvassing for 2010 one year earlier, in 2009, sending enumerators to every block in most of the country to verify or correct address list entries; this was the one 2010 census operation that was able to make use of handheld computers. Looking ahead to 2020, the Census Bureau has launched its Geographic Support Systems Initiative, which I and my panel colleagues have followed with great interest and which we endorsed in broad outlines in our first interim report.

As this work progresses, I make the following suggestions:

- First, *the Bureau should not enter into its geographic research with a preconceived notion that either MAF or TIGER is an unassailable “gold standard.”* To be clear, I am not suggesting that the Bureau is currently laboring under such a notion; indeed, I think that the Bureau has been quite candid in noting shortcomings and in suggesting the need for quality metrics. This is just to say that change and improvement are only possible if it is acknowledged that there is room for improvement.
- There are significant limitations at present to the use of purely commercial mapping resources (e.g., Google Maps, Nokia/Navteq, databases maintained by ESRI) in the census context, or even the address list resources of the U.S. Postal Service). What I think is important is that *the Bureau’s geographic research should focus on the coverage properties of MAF/TIGER and those of alternative resources.* The census has to accurately represent all subgroups of the nation’s population, and some of these groups live in locations or housing stock where standard addressing procedures are lacking and where field enumeration practices can be extremely challenging; this is as true of populations in intensely remote locations (e.g., Alaska Natives) as it is of residents of intricately built-up locations (e.g., individual blocks in New York City). Very little is actually known about the quality of spatial data or its impact on the accuracy of a census, or the subsequent demographic estimates that derive from those data.

- Finding the right balance between field work (direct address canvassing or “map spotting”) and drawing from existing computerized resources will be difficult; my own sense is that neither of the “pure” outcomes of zero canvassing or 100 percent canvassing is going to be satisfactory. So, I reiterate my hope: *the Census Bureau should consider the same kind of “adaptive”/responsive/flexible approach for updating its geographic resources as it hopes to implement in its field operations.* That is, by considering the coverage properties of various alternative data sources for some subgroups or geography types and balancing those with the cost/quality trade-offs associated with new field data collection methods, I hope that the Bureau can avoid the same “one size fits all” approach that has driven census operations in the past. The geography and housing/address stock of downtown Chicago is different from that in pueblos in New Mexico, and the best means of ensuring up-to-date geographic coverage in those areas will vary, too.
- *The ability to effectively plan is predicated on the ability to commit Bureau resources into prospective testing of alternative field data collection methods and to appropriately assess the impact of alternative methods on coverage of addresses.* Retrospective testing (looking at past patterns of coverage and the operations that produce it, for example) can be a powerful way to address questions about how operational procedures relate to Census coverage and to model the trade-offs between costs for data collection and anticipated coverage. However, only prospective testing provides the promise of assessing alternatives that have not been previously considered.
- In the maintenance of its geographic resources as in the reshaping of its field operations, the panel noted in its first interim report that *the Census Bureau can learn a great deal from outside the Bureau—from private and public sector organizations faced with similar challenges as well as from statistical agencies in other countries.* That is, the Bureau should consider the techniques used by commercial map vendors in updating their products, draw from the experience of firms such as UPS, and study the specific operations conducted by agencies like Statistics Canada.

In closing, I understand that recent developments involving the American Community Survey (ACS) are a secondary topic for today’s hearing. As a regular user of ACS data for a wide range of projects, I would be remiss if I did not take this opportunity to state my own personal hope that the Senate will undo the appropriations amendments passed in the House to make ACS response voluntary rather than mandatory and then to cut off funding to the ACS altogether. And I would welcome the chance to answer any questions regarding uses of the ACS. But—in keeping with the main theme of the hearing—I would like to close by stating my belief that a healthy, vigorous ACS is critical to an improved 2020 census, and essential to a worthwhile and effective planning process for that 2020 census. Of the Census Bureau’s other major activities, the ACS is a particularly strong test-bed for 2020 census approaches and systems. ACS field operations include a nonresponse follow-up component that permits address list updating by field data collection staff, particularly in rural areas where the regular MAF updates (from the U.S. Postal Service) are thought to be weakest; hence, the ACS is an ideal forum for testing geographically-targeted updating of the address list base to areas of suspected

undercoverage. The “methods panel” in the ACS gives the Bureau an opportunity to pilot-test revised questionnaire wording or formatting of concepts. And, perhaps most fundamentally, implementation of an “adaptive” data collection process in the ACS is not just a useful test-bed but a critical proving ground, before such process rolls out in the 2020 census. Utilization of ACS as a test-bed is not a new idea—Director Groves and many other outside observers have suggested precisely that—but I think it is a critical one for decision-making aimed at optimizing coverage in light of cost constraints. Of course, the ACS cannot function as a test-bed for 2020 if it does not exist (or exists in a severely hobbled form) after enactment of fiscal year 2013 appropriations, and so (again, my personal opinion) I urge the Senate to support continuance of the ACS.

I thank you again for the opportunity to testify before you today and I welcome your questions.