

Discussion

*Brad Edwards*¹

Nancy Bates has pioneered an important subfield in survey methods: hard-to-survey populations. The topic of the Bates article in this issue is the U.S. Census' experience using social marketing to improve mail and web response rates for these populations. Populations can be hard to survey in many ways. The article is in line with a whole body of work Bates has developed on methods for asking about race, Hispanic origin, sexual orientation, and other characteristics, with a major focus on survey and census nonresponse. I was privileged to work with her on what became a major milestone in this field, an international conference in 2012 on hard-to-reach populations, and on the book that followed (Tourangeau et al. 2014). The conference was also the springboard for a special issue of the *Journal of Official Statistics* on the hard-to-reach (Willis et al. 2014).

The U.S. Decennial Census is a unique vehicle for studying the role social marketing can play in improving response from hard-to-survey groups because of its very large scale and its mandate to count every single person living in the United States. Other large scale vehicles, such as the U.S. presidential elections or national campaigns to change health behaviors, do not have such an exacting goal or such broad scope. Sample surveys are much smaller in scope and lack the resources to explore social marketing. The capability to target small areas like census tracts or even individuals in a sample at very low cost has only just emerged.

I work on surveys, not censuses, and for a contracting organization, not the U.S. Census. Most of my work has been on face-to-face surveys, not mail or mixed-mode like the Decennial Census. Nonetheless, there are many parallels in our work, and the program at the Census may become a model that brings social marketing into the forefront of survey methods.

The genesis of the research summarized in the Bates article comes from the observation that it is getting harder for censuses in the developed world to count a nation's people. There are many reasons for this, but changes in the concept of privacy is a prominent one. In the developed world, very few of us can live "off the grid". Countless electronic transactions mark our daily lives, and conventional notions of privacy have become outmoded. Almost everything about us seems public, but the notion persists that our personal data belong to us. It is "our" data. With that comes the recognition that our data is worth something to organizations, that we should get something for it, something like a discount at the supermarket, something greater than the satisfaction of contributing to the public good. It is not clear to many Americans anymore what they get from surveys and

¹ Westat, Inc., 1600 Research Boulevard Rockville, MD 20850-3195 Maryland 20009, U.S.A. Email: bradedwards@westat.com

censuses; neither is it clear what happens to it. When we give our data to organization A, ownership of the data is shared between us and A. We can then share our data with organization B, but A can also choose to share the data with another party, C, unless we have expressly forbidden it. It is easy to see how control – “ownership” – can be lost.

The late Eleanor Singer studied the relationship between privacy and willingness to participate in surveys. Her research showed that for most people, the decision to participate is the result of a cost-benefit tradeoff: What’s in it for me? (Singer 2016, 2011, 2003). The length of time, the cognitive burden, the topic salience, the cash or non-cash incentive, what I’ve heard about it, what my friends and neighbors are likely to think about it, the interviewer’s charm – all of those may be taken into account. But society’s benefit, the common good, is less likely to enter into the equation than in the past. Trust in government, and in institutions like the press, have plummeted over the past decades.

Survey methodologists can manipulate some of these factors, of course (better publicity and advance letters, shorter interviews, larger incentives, interviewer training), but at a cost. Alongside response rate declines, we see costs increase. Many repeating surveys have spent increasing amounts of money to slow the response rate slide. (The Decennial Census from 1990 through 2010 is a prime example.) Because of small sample sizes in local areas, geographically targeted advertising has held little value for surveys: costs would be prohibitive, and advertising in a small geographic area could increase disclosure risk because it would be an indicator of the specific small second stage sampling units. What is so appealing about social marketing is that messages can be *tailored to individuals at very low cost*. Ideally the message prompts the individual to take action, and makes it very easy to act (by clicking a link in the message, for example, to reach the online census form).

The Census test of digital advertising in Savannah was a big leap down this road. It established that digital ads could drive hard-to-survey groups to online forms, increasing response rates and decreasing cost. This is the Holy Grail for survey methodologists. One can imagine many studies springing from it. Randomized control studies could test the effectiveness of different digital ads for different hard-to-survey groups; ad awareness and recall studies could explore the role that message exposure plays in the survey response decision; advertising metrics (reach, click-through rate, recall lift) could support comparative research on cost effectiveness of ad buys.

Underpinning the social advertising program at Census is the Low Response Score (LRS). It builds on the Hard to Survey (HTS) metric that Census developed for 2010. These response propensity models have stirred great interest for the survey world. Other organizations have been keen to apply them to gain efficiency in deploying resources, targeting ad campaigns and outreach materials, staging data collection efforts, and making weighting adjustments. The potential usefulness of the LRS in adaptive design is clear. It could help determine when phase capacity is likely to be reached, and when to change to a different approach. It could allow precise tailoring of each phase down to the block group or even the individual level for different population groups (personas, to use the Bates term).

Researchers at UCLA and Westat sought to use the HTS model as one of many data sources to explore the possibility of nonresponse bias on the California Health Interview Survey (CHIS), an annual telephone survey conducted for the State of California. Unexpectedly, no relationship was found between the achieved CHIS response rates and the HTS. (Lee et al. 2009) We speculated that, because the HTS was based on mail

response, it incorporated factors that were unique to mail, and did not include factors that may be unique to other modes. This speculation seemed to be supported by a more recent Westat analysis of the HTS and the LRS compared to achieved response rates on a large face-to-face data collection: again, no relationship. In 2017 Westat researchers completed a comparison of response rates on a *mail* survey with the HTS and the LRS, and again, found no relationship, concluding that those tools offer very little information to predict mail screener return on their study. Perhaps the mandatory nature of the Census invokes factors in propensity to respond that are not present in other mail surveys. We are currently building our own model for response to the face-to-face mode, based on our experience across a number of large household surveys. For face-to-face, gated communities and locked buildings are correlated with higher nonresponse rates. The impact of linguistically isolated communities is arguably also larger in face-to-face, because they are hard to identify and it is difficult to deploy interviewers on the ground who speak the language.

Bates' work is rooted in audience segmentation and social marketing. Running through this is the concept of social distance. The Census is trying to reach out to hard-to-survey audiences in ways that reduce the social distance, speaking to them in their language if you will. As the nearly two billion Facebook users know, suggestions from friends are very powerful, even if they are only "Facebook friends". The social media platform paradigm is built on "likes." It is an extension of an axiom from advertising: you will have more success if you begin with what the customer needs, not what you have to sell. This is the opposite of the survey paradigm, which start with the researcher's question, "What do I want to know?"

I was especially struck by the "cynical fifth" mindset in the Census segmentation. People with this mindset can have quite negative attitudes about the census or surveys or government, saying things like "It never does me any good," yet they may hold a basic belief that it *should* do something good, that it would be fair if only it could do something good, whatever that might be. This suggests the possibility of some common ground, where an ad campaign or an interviewer could turn the attitude around. For example, in an earlier draft Bates noted that many with this mindset have a strong belief in the U.S. Constitution; branding the census as something written into the Constitution might persuade them to respond.

A dark side travels with the rosy promise of social marketing for censuses and surveys, however. The success of social marketing depends on vast amounts of information being readily available, which strains the security apparatus to prevent unauthorized disclosure. One of the largest private data breaches ever detected was disclosed this summer. Personal identifying information for 143,000,000 adults was hacked from Equifax, one of three organizations that provide information about credit worthiness of individuals to lenders in the U.S. (White 2017). The potential for harm is great. The hacker could find a ready market on the dark web for these data, and its value could be retained for years into the future.

Planning for the 2020 decennial is in a critical stage. In the previous three decades, spending for the decennial has ramped up dramatically in the second year before the census, to levels two or three times the level of spending in the earlier years. At this writing, the government is operating on a continuing resolution, holding spending to the same level as last year, and Congress will be debating funding levels for Fiscal Year 2018

in the fall of 2017. With a requirement to spend less on the 2020 decennial than in 2010, the Census has embarked on an ambitious program of modernization, including a much more sophisticated social marketing campaign. The program requires an investment in new technologies and analytics. Without the funds now to ramp up and adequately test the program, the Census may have to fall back on more traditional methods, run out of funds, and risk missing millions of Americans who are hard to survey. A cynic might wish for that to happen, if those missed would be unlikely to vote in the cynic's favor. But that would torpedo one of the basic underpinnings of America's democracy, the principle of one person-one vote. Without an accurate count of persons, that principle is lost.

The U.S. Census has shown tremendous resilience over the past century, navigating some treacherous partisan waters and developing innovative methods that transform our field. It has never had a breach, and has never shared with anyone the decennial data it must guard for 70 years. It has a good brand. We can all hope that science and the public's common sense will prevail in these troubled times. But a stable budget and a supportive administration would help a lot. Better yet, move the Decennial Census out of the annual discretionary budget process, and onto a 10-year funding cycle.

References

- Lee, S., E.R. Brown, D. Grant, T.R. Belin, and J.M. Brick. 2009. "Exploring Nonresponse Bias in a Health Survey Using Neighborhood Characteristics." *American Journal of Public Health* 99: 1811–1817. Available at: <http://ajph.aphapublications.org/doi/pdf/10.2105/AJPH.2008.154161> (accessed 13 October 2017).
- Singer, E. 2003. "Exploring the Meaning of Consent: Participation in Research and Beliefs about Risks and Benefits." *Journal of Official Statistics* 19: 273–286.
- Singer, E. 2011. "Toward a Benefit-Cost Theory of Survey Participation." *Journal of Official Statistics* 27: 379–398.
- Singer, E. 2016. "Reflections on Surveys' Past and Future." *Journal of Survey Statistics and Methodology* 4: 463–475. Doi: <http://dx.doi.org/10.1093/jssam/smw026>.
- Tourangeau, R., B. Edwards, T.P. Johnson, K.M. Wolter, and N. Bates. 2014. *Hard-to-Survey Populations*. Cambridge University Press.
- White, G.B. 2017. "A Cybersecurity Breach at Equifax Left Pretty Much Everyone's Financial Data Vulnerable." *The Atlantic*, September 7, 2017. Available at: <https://www.theatlantic.com/business/archive/2017/09/equifax-cybersecurity-breach/539178/> (accessed October 13, 2017).
- Willis, G.B., T.W. Smith, S. Shariff-Marco, et al. 2014. "Overview of the Special Issue on Surveying the Hard-to-Reach." *Journal of Official Statistics* 30: 171–176. Doi: <http://dx.doi.org/10.2478/jos-2014-0011>.

Received October 2017