

Save your census

National censuses and surveys are threatened around the world by high costs and low response rates. The demographic data they yield are too valuable to lose, warn **Stephen E. Fienberg** and **Kenneth Prewitt**.

Census-taking around the world is under assault, thanks to concerns about privacy, cost and response rates. Most scientists and policy-makers worldwide fail to appreciate what is at stake until it is too late to repair the damage of short-sighted decisions. In June, Canada's government announced that it will turn its mandatory long-form questionnaire, sent to 20% of the population, into a voluntary survey sent to 30%. This decision will lower the quality and raise the cost of information on nearly every issue before Canada's government.

Government statistics are no less vital to a nation's scientific infrastructure than is an observatory or a particle accelerator, and need stable funding and protection. Detailed, reliable demographic data are used in a vast array of policy decisions and research studies, from determining how many hospitals are needed to tracking whether the ongoing poverty of a group can be linked to health or education. Census data provide the gold standard against which all other studies on such issues can be corrected and judged.

National statistics agencies must learn to make use of administrative records and digital data, but much work is needed to ensure that this approach will produce data that are equivalent or superior to those produced by current methods.

Apathy and expense

As the 2010 US census got under way, privacy was a major issue, highlighted by conservative talk-show hosts and even a member of Congress, Michele Bachmann (Republican, Minnesota), who opined that questions related to age, gender and ethnicity were too invasive. 'Privacy' concerns make for good sound bites, but the fact is that no one in government is more zealous about privacy protection than national statistics officers.

There are more serious problems. One is cost. The 2010 US decennial census cost about US\$13 billion, double what was spent on the 2000 census, which in turn doubled the 1990 cost. The reasons include population growth, more mobile household members and increasing language barriers. Changing to a voluntary survey is not the solution — this increases costs, by requiring that more forms be sent out to get the same response rate, and produces lower-quality data.



People are less responsive to surveys than they were in 1930.

A second concern is that, in increasing numbers, people can't be bothered to respond to information requests. Telemarketing has soured the environment for phone interviews, and junk mail clutters e-mail inboxes. Who has time to distinguish legitimate surveys from the flood of lookalikes? Item non-response is reaching unprecedented levels: from the 1990 to 2000 US censuses, the number of people who skipped a question on monthly rent, for example, rose from 3.3% of the population to 15.6%. Reputable academic surveys struggle to get a 50% response rate to entire questionnaires, compared with common rates of 70% two decades ago. Government agencies hold themselves to much higher standards. But it is hard going. This is all the more reason for making census responses mandatory.

Critics of the traditional mail-out census often point to other countries as examples of how things can be done differently. Nations including Sweden, the Netherlands and Denmark switched from censuses to systems based on data such as national registrations, tax records and parish books, mostly in the 1990s, mainly for reasons of cost. But administrative records are often incomplete, inaccurate and out-of-date. Some nations have regretted their move. Germany, for example, believed that it had good registers when it adopted this system in 1987, but belatedly learned that the records fail to capture the movement of people across borders within Europe. A combination of administrative

records, from post-office address lists, driving-licence records and health registers, has been proposed to replace the UK census after 2011. This will be difficult. UK statisticians have a decade to try to work out problems with this methodology before their next effort in 2021.

Official administrative records are just the tip of the iceberg. Each person leaves a digital footprint when swiping a bank card, surfing the Internet, getting on a plane or using Global Positioning System (GPS) tools to navigate. Governments will inevitably ask whether these sources can provide a cheaper, quicker stream of data about a nation's population than censuses or surveys.

Such data are potentially powerful. Urban planners could, for example, use GPS data from mobile phones to help determine rush-hour travel patterns. But there are serious problems with these mostly proprietary sources. There needs to be protection of personal details — something which commercial companies such as Facebook have been heavily criticized for failing to provide. There need to be quality standards: data useful to the marketing departments of pharmaceutical or telecom companies aren't good enough for public policy choices. Above all, data should be representative. Digital data

reflect those who use a technology, not necessarily the wider population.

The task ahead is formidable. Census communities need to gather ideas for creating more robust data

collection using administrative and electronic sources to supplement the long-form surveys and to help address problems of non-response. National census offices need to become more nimble in testing methodologies without jettisoning the best of current approaches. Scientists need to appreciate that a vital scientific resource is on the line. ■

Stephen E. Fienberg is in the departments of statistics and machine learning at Carnegie Mellon University, Pittsburgh, Pennsylvania 15213-3890, USA. **Kenneth Prewitt** is in the School of International and Public Affairs, Columbia University, New York, New York 10027, USA. e-mail: fienberg@stat.cmu.edu

"Demographic data are used in a vast array of policy decisions and research studies."