2B or Not 2B? What Should Have Happened with the Canadian Long Form Census? What Should Happen Now?
Michael R. Veall

Canadian Public Policy, Volume 36, Number 3, Septembre/septembre 2010, pp. 395-399 (Article)

Published by University of Toronto Press
DOI: 10.1353/cpp.2010.0007

For additional information about this article
http://muse.jhu.edu/journals/cpp/summary/v036/36.3.veall.html
2B or Not 2B? What Should Have Happened with the Canadian Long Form Census? What Should Happen Now?

MICHAEL R. VEALL
Department of Economics, McMaster University
Hamilton, Ontario

Since Confederation, Canada has had a mandatory census with questions on religion, ethnic background, occupation, school attendance, literacy, infirmities, and much else (see Dillon 2010). Many of these questions are now on the long form census (2B) that since 1971 has gone to 20 percent of households. The remaining 80 percent of the households received the short form census (2A), which has basic questions on age, sex, marital status, family relationships, and mother tongue. In late June 2010, six weeks before the forms were to go to the printer, the federal government of Canada announced without public consultation that in 2011 the census would consist of only the short form, which would go to all households. The previously mandatory long form would be eliminated and replaced by the voluntary National Household Survey (NHS).

The government has argued that the data will continue to be of high quality with less state coercion. Yet there is insufficient evidence to be confident about the data quality, and substantial risk that the NHS will be a costly failure. Even given the government position that a mandatory survey is overly intrusive, there should have been advanced testing of the voluntary survey approach. I discuss some limited options that might mitigate the risk going forward with a purely voluntary survey.

WHAT IS THE RISK?

The risk of the voluntary approach is that the non-response bias may be high. The people who respond may be different from those who do not. The mandatory Canada Census has a minimal non-response bias because the response rate is so high. But even an overall voluntary response rate of 80 to 85 percent (much higher than Statistics Canada expectations for the NHS of 65 to 70 percent) could admit substantial non-response bias for some census tracts and some groups, such as immigrants or those with low income, low formal education, or neither English nor French as a first language. Non-response bias is particularly difficult because it is invisible; it depends upon the characteristics of non-respondents, which the statistician of course does not observe. Any loss of accuracy and comprehensiveness in the 2011 data would also entail the loss of comparability across time periods and with the data of other countries (see Dillon 2010).

How have the government and its supporters tried to defend the position that the new data would be just as high quality?

The first defence was that the sample of responses would be at least as large because the number of
forms distributed would be increased from one-fifth of the population for the mandatory long form to one-third for the NHS to allow for increased non-response. However, this approach will not prevent non-response bias. As Green and Milligan (2010) consider this important aspect, I will not discuss it further here except to note that the Industry minister appeared to suggest that the problem could be solved using non-response weights drawn from the mandatory short form. But this cannot counter the problem of response bias associated with, say, those of low formal education, because the short form does not have such data. In addition, a major advantage of mandatory census data is that there is little reliance on the assumption that those who do respond within a subgroup are representative of those who do not, an assumption that is required for non-response weighting to be effective. (For a more detailed discussion, again see Green and Milligan 2010).

The second defence was that the mandatory approach increases the probability of lying by respondents. This does seem likely, in part because those who might lie would simply not respond in a voluntary setting. The question is how large this problem is compared to the non-response problem. My view is that the government should have prepared evidence on this question to bolster its position and to help preserve comparability of census data over time. This is something that could easily have been done.

In the case of income data where mandatory census data can be checked against personal income tax data (see, for example, Frenette, Green, and Picot 2004), the evidence suggests that lying is not an important contaminant of the data.

While there were other arguments involving intrusiveness and other countries moving away from the census approach, the only other defence of prospective NHS data quality that I know of referred to the failed experiment with voluntary response run by the American Community Survey (ACS). The summary reports (US Census Bureau 2003; 2004) maintained that the voluntary approach could have worked with more resources. However, this is speculative, given that the experiment did not actually work. Second, the ACS is not a census but rather a smaller monthly survey that replaces the US long form and bridges the gap between US decennial censuses. Unlike the one-shot NHS, a problem one month can be fixed the next, and survey employees can learn over time. Third, the reports indicated that telephone and mail response rates for blacks and Hispanics were markedly lower with the voluntary approach and appear to maintain that this problem could in part be fixed by non-response weights, apparently to come from a still mandatory US decennial census. This option would not be available to the NHS.

**WHAT SHOULD HAVE HAPPENED**

Suppose that after conducting the 2006 Census, the government had stated it had concerns about intrusiveness, announced that it was considering making the long form voluntary for 2011, and embarked upon a public consultation. This process could have replaced the actual Statistics Canada consultation process (which in retrospect was less valuable than it might have been, given that it was conducted under the presumption that the long form would remain mandatory). There would no doubt have been the same clash witnessed in recent months, with some seeing the state as having no right to such information and others seeing mandatory participation as akin to jury duty or paying taxes as a requirement for the functioning of government. But I believe that some would have asked, “How will the quality of the data from the voluntary approach compare to the quality of the previous data?”

In my view, this question should have been answered with a pilot study. A sample could have been drawn from those who did not complete the mandatory long form in 2006 and this group could have been given the survey on a voluntary basis. These individuals could have been asked in addition...
to comment on the intrusiveness of the form item by item, providing information on question selection. If it was too late to test the 2006 Census, test runs could still have been done based on the mandatory Labour Force Survey, which has some comparable questions. A technical study could have been prepared and made public that compared the data from mandatory and voluntary approaches to administrative data such as personal income tax data.

What might the conclusions have been? If it were found that the voluntary approach did very well, it could have been adopted. If the voluntary approach did reasonably well, with some shortcomings, the side-by-side trial could have yielded correction factors to be applied to subsequent voluntary surveys and could have preserved some comparability with census information. If the voluntary approach did poorly, the government would have had either to accept a mandatory approach or not run a survey at all. Decisions regarding the census could have been based on evidence.\footnote{14}

WHAT SHOULD HAPPEN NOW?

One possible way forward for the government is to move toward the proposal of its own National Statistics Council\footnote{15}, running the 2011 Census with a mandatory long form while trying to lay the groundwork for running a less intrusive and perhaps voluntary long form for 2016. This would be my personal preference. But what steps might be taken to improve data quality if the government maintains its position regarding a voluntary NHS for 2011?

1. Those who choose not to respond to the NHS could be required to sign off on a written form or by Internet. This would provide a mechanism to verify that as many as possible have had the opportunity to respond and have paid at least minimal attention to the request.

2. While, as I have argued, non-response weighting is far from a panacea, it is possible that useful NHS non-response weights for income could be drawn from personal income tax data. More generally, access to suitably anonymized personal income tax data could be made similar to current access to census data, in particular at the census tract level.

3. As the Labour Force Survey is still mandatory, it would still be possible to have a mandatory versus voluntary test based on it. Indeed, if the government is arguing that mandatory surveys are inherently inaccurate because of lying, such a test would seem a priority, given the importance attached to such measures as the unemployment rate.

CONCLUSION

In defending its June 2010 decision to eliminate the mandatory long form census and replace it with a voluntary survey, the government argued that the new data would remain of high quality. However, there is insufficient evidence to support this claim. A pilot study could have provided the basis for more accurate, comprehensive data with better understood limitations, even given the government insistence upon a voluntary survey.

If the government does not take the advice of the National Statistics Council and a voluntary survey is used in 2011, some steps may be taken to lower the risk of failure, including a mechanism to ensure that all those chosen to receive the voluntary survey acknowledge that they have had an opportunity to respond, and the possibility of using personal income tax data to estimate non-response weights. The wider availability of tax and other administrative data is also advocated. However, these suggestions do not constitute a complete solution nor, since they have not been tested, can they be advanced with complete confidence.

NOTES

I thank Tom Crossley, Don Drummond, Deb Fretz, John Helliwell, Kris Inwood, Kevin Milligan, Doug Norris, Arthur Sweetman, and an individual within Statistics.

\* \* \*
Canada who chooses to remain anonymous for related discussions. All opinions are mine.

1 “Sen. Marjorie LeBreton, who leads the Tories in the Senate, suggested nothing will be lost by moving to a voluntary long census: ‘There’s still going to be a long form,’ she said in Ottawa. ‘The only difference is, this is voluntary. Canadians, I believe and we believe, will be very happy to fill in the long forms,’ she said” (Canada Press, 13 July 2010, at http://www.ctv.ca/CTVNews/Canada/20100713/long-form-census-scrapped-100713/). From an interview with Industry Minister Tony Clement, Globe and Mail, 21 July 2010: “Q: do understand your rationale. What I don’t understand is this: the impression we’ve got from your comments over the last few days is that Statscan is A-OK with this. That this is a perfectly acceptable substitute for the mandatory long form—and they signed off on that. That they said this will satisfy everything necessary to produce a statistically valid census that is no different from the last one. A: Right and I do assert that. When an agency of government reports to its minister and gives that minister options, I am entitled to assume that they are comfortable with those options” (at http://www.theglobeandmail.com/news/politics/ottawa-notebook/tony-clement-cleans-the-air-on-census/article1647055/). And in the Globe and Mail, 5 August 2010: “Mr. Clement remains adamant that changes to the census will not affect the quality of data collected, even though statisticians warn a voluntary system will not be as accurate … ‘All these provincial governments and these social institutions and private businesses—we’ll get them some data that will be useful and reliable,’ he said” (at http://www.theglobeandmail.com/news/politics/census-critics-just-want-easy-ride-clement-says/article1663469/).

2 Mr Clement told CBC Radio on 14 August 2010, “I think we are meeting the objections of those who are concerned about robust and reliable data” (at http://www.cbc.ca/politics/story/2010/08/13/clement-census-interview.html). All web pages accessed 17 August 2010.

3 There will be new expenditures of $30 million on advertising and costs associated with the larger NHS sample. See, for example, http://www.ctv.ca/CTVNews/TopStories/20100720/census-clement-100720/, accessed 16 August 2010.


5 The industry minister appeared to acknowledge that a mandatory approach was more accurate when asked in an interview by CBC Radio (17 July 2010) why the Census of Agriculture would remain mandatory. He responded, “The argument obviously to farming associations and to farmers is, ‘You fill out the form, it'll help the government help you in your farming activities’” (accessed 15 August 2010 at http://www.cbc.ca/politics/story/2010/07/16/lement-house-census.html). It is difficult to understand why that argument applies just to the agricultural census and farmers.


7 The only example given that I know of was that of 20,000 Canadians responding that they were Jedi Knights in the 2001 Census. See the email from the Prime Minister’s Director of Communications reproduced at http://www.cbc.ca/politics/insidepolitics/2010/07/dimitriwatch-but-do-jedi-knights-support-the-mandatory-census-long-form-or-not.html. Note, however, that this is a tiny percentage of respondents (less than 0.1 percent), and it is not clear that the number would have been lessened if Canada had had a voluntary survey. Of course, the total percentage of false responses could be higher for this and other questions.

8 The Frenette, Green, and Picot study was not, however, designed to study the question of lying, and hence its results are only suggestive. On the other hand, the lying issue has been considered by Statistics Canada, which uses a number of techniques to mitigate the effects in data aggregates. The Statistics Canada methods of imputation and editing (the Canadian Census Imputation and Editing System, or CANCIS) are used around the world.

9 Not all the intrusiveness defences seemed well founded. Government criticisms of particular questions, penalties, or the conduct of the census were criticisms of matters it itself controlled. A case in point was the use by the Industry minister of a “what time do you leave for work” question as an example of excessive intrusiveness when that question was not in the 2006 Census but had been newly approved by cabinet for the NHS. See commentary by John Geddes of Maclean’s, accessed 16 August

10 See the 21 July 2010 blog of Tim Powers, a former director of policy and research for the leader of the Progressive Conservative Party of Canada, at http://www.theglobeandmail.com/news/politics/second-reading/silver-powers/come-to-your-senses-on-the-census-please/article1647521/, or Neil Reynolds’s Globe and Mail column of 26 July 2010 at http://www.theglobeandmail.com/news/opinions/the-anachronistic-coercive-unnecessary-census/article1650049/. The United Kingdom was one example given, but it will be having a census in 2011. There is discussion about not having it in 2021, providing a significant period for additional consultation (see Thompson 2010). Another example given was Denmark, which avoids a census by having a registry system, connecting all government records by an individual ID number. Most would argue that is a far more intrusive system than a mandatory census, although partial implementation would be possible. The Danish system took about ten years to develop. See the Globe and Mail, 26 July 2010, at http://www.theglobeandmail.com/news/politics/european-census-alternatives-have-privacy-concerns-of-their-own/article1652595/ All web pages were accessed 16 August 2010.


12 The US decennial census has questions on race and Hispanic status, permitting non-response weighting on those characteristics. It does not have questions on income or education, and hence non-response weighting would not be possible on those characteristics.

13 Something like this was done in the 2008 and 2009 test censuses, which were voluntary (see Parent and Lathe 2009). I know of no published results from the 2009 test census. For the 2008 test census (which was somewhat stratified), Parent and Lathe report a raw response rate of 46 percent on the long form. However, the emphasis of the exercise was not on encouraging response rates (especially as at the time it was assumed that the mandatory long form would remain in 2011) but rather on content and other technical comparisons.

14 Statistics Canada would also have gained valuable technical information about how to run such a survey.


REFERENCES


