Component of Statistics Canada Catalogue no. 11-522-X
Statistics Canada's International Symposium Series: Proceedings

Article

Symposium 2008:
Data Collection: Challenges, Achievements and New Directions

Mailing Census Questionnaires: Results from 2006 and Plans for 2011

by Sander Post

2009

Statistics Canada Statistique Canada

Canada

Proceedings of Statistics Canada Symposium 2008
Data Collection: Challenges, Achievements and New Directions

Mailing Census Questionnaires: Results from 2006 and Plans for 2011

Sander Post[1]

Abstract

Prior to 2006, the Canadian Census of Population relied on field staff to deliver questionnaires to all dwellings in Canada. For the 2006 Census, an address frame was created to cover almost 70% of dwellings in Canada, and these questionnaires were delivered by Canada Post. For the 2011 Census, Statistics Canada aims to expand this frame further, with a target of delivering questionnaires by mail to between 80% and 85% of dwellings. Mailing questionnaires for the Census raises a number of issues, among them: ensuring returned questionnaires are counted in the right area, creating an up to date address frame that includes all new growth, and determining which areas are unsuitable for having questionnaires delivered by mail. Changes to the address frame update procedures for 2011, most notably the decision to use purely administrative data as the frame wherever possible and conduct field update exercises only where deemed necessary, provide a new set of challenges for the 2011 Census.

Key Words: Questionnaire delivery, Census, Mailing.

1. Introduction

Many aspects of the data collection process for the 2006 Census were significantly updated from the data collection process used for the 2001 Census. These changes were largely enabled by significant use of information technology (IT) infrastructure and included giving respondents the opportunity to respond via the internet, centralizing follow-up activities, and reducing the number of and increasing the specialization of field staff.

One of the significant changes concerns the basic task of getting the questionnaire to the dwelling. Prior to 2006, all questionnaires were hand delivered by Census staff. For the 2006 Census, a majority of questionnaires were delivered via the Canada Post mail delivery system. Changing questionnaire delivery provides many benefits to the Census, but also posed many challenges. This paper will describe both the benefits and challenges.

2. Census questionnaire delivery prior to 2006

2.1 Enumeration

In the 2001 Census, Canada was divided into 42,851 Enumeration Areas (EA). Each EA consisted of one or more blocks, and typically contained around three hundred dwellings. On or around Census Day, field staff dropped off questionnaires at each dwelling in their EA. At the same time, they listed each dwelling and classified the dwelling as vacant or occupied. They subsequently edited each questionnaire and followed up if the questionnaire failed the edits, and followed-up on each dwelling that hadn't responded by a specific date.

While there are some minor variations, the above description applies to the collection process for all Censuses from 1971 to 2001. This approach was both effective and successful in its time, though very costly. However, there was clearly a potential to make the process more efficient, and to address the challenges of hiring large numbers of staff as well as the challenge of confidentiality posed by the use of local enumerators. This potential was certainly a result of the growing scope and power of Statistics Canada's IT infrastructure.

[1]Sander Post, Statistics Canada, 150 Tunney's Pasture Driveway, Ottawa, Ontario, Canada, K1A 0T6, sander.post@statcan.gc.ca

Prior to 2006, each Census required that field staff list addresses for all dwellings in Canada. The use of the Address Register and the Spatial Data Infrastructure made much of this effort unnecessary. The Spatial Data Infrastructure (SDI) is a spatial database containing road network, legal boundaries, and Statistics Canada boundaries for all of Canada. The Address Register (AR) is a database of addresses, primarily built up from the previous Census and linked to previous Census geography, and updated intercensally from administrative sources. The SDI contains road names and address ranges, allowing AR addresses to be linked to a unique blockface.

2.2 Issues

One key aspect of 2001 collection operations the Census wished to improve upon was the size of the work force. The 42,851 Enumeration Areas used in 2001 required that the hiring and training of a temporary work force of more than 40,000 people to handle all collection tasks. Hiring and managing a temporary workforce of this size is a complex and costly activity and any opportunity to reduce the workforce required, to increase the specialization of employees, or to spread out the work over a longer time frame would make the process simpler and more efficient.

One key lesson learned from coverage improvement operations performed in 2001 using the Address Register is that the Address Register is, for the more urban parts of the country, a very good address frame. Given a complete and up-to-date address frame, questionnaires can be sent through the mail instead of being hand delivered. In areas where Canada Post delivers the questionnaires, Statistics Canada will no longer needs to conduct the labour intensive operation of listing and delivery around Census Day.

However, Statistics Canada also requires that the address frame be linked to Census geographies in order to associate the dwelling and its occupants' characteristics to the correct location. In 2001, this link was established when field staff dropped off a questionnaire, as they also recorded which block the dwelling was in. When the questionnaire was returned to Statistics Canada, the dwelling information would be tabulated in the right block. An address frame for mailout also needs to be linked to Census geography, so that returned questionnaires can be tabulated for the right block.

Finally, there must be methods to up-date and maintain both the address and geographic frames in areas of high population change during the last few months leading up to Census Day. The most complete way of verifying this information is to conduct field operations (listing) that visit all areas of the mailout and correct and update both the address and geographic data. For data collection operations in 2006, the Census wished to minimize the field staffing requirements and improve efficiency by both delivering questionnaires by mail and by doing the needed listing work for mailing before Census Day.

3. Questionnaire delivery in the 2006 Census

3.1 Defining mailout

The Census has well defined objectives, and the change in the method of questionnaire delivery cannot affect these objectives. The Census needs to count the whole in-scope population of both residents of Canada and the housing stock, and count them in the right block, Statistics Canada's lowest level of dissemination geography.

Canada Post's mandate is to provide mail service to all dwellings in Canada. To deliver to some dwellings, the letter carrier must know the name of the occupant, as the dwelling itself has no address, or if it does, it is not used by Canada Post in their delivery system. As the Address Register does not maintain occupant name against the address, such dwellings cannot be mailed to, and must continue to have questionnaires delivered by Census field staff. The only information known before questionnaire delivery about a dwelling is its civic address. It follows that the Census can mail to all areas of the country where Canada Post provides delivery based on the civic address.

3.1.1 Required characteristics of the frame

To meet Census requirements within a mailout area, an address frame (mailing list) requires the following characteristics:

1. It represents a well defined spatial extent consisting of complete blocks. Collection units are made of blocks, and blocks are required to be entirely within mailout or entirely excluded from mailout.
2. Within this spatial extent, Canada Post must support address-based delivery.
3. The Census must have addresses for all dwellings within the mailout area.
4. All addresses must be linked to a block, for tabulation/dissemination purposes.
5. Virtually all addresses should have a postal code.

With regards to this fifth requirement, postal codes exist for all dwellings, but cannot always be determined by our automated systems. However, in order to obtain the best postal rates for a mailout of this magnitude, Canada Post requires a very high rate of postal coding, as the presence of a postal code greatly facilitates their internal operations.

In addition, under the assumption that the mailing list will be completely field verified, the following characteristics are also required for ease of field operations:

6. Contiguous mailout areas must be above a specific size threshold (about 3000 dwellings) and must not contain any small pockets of non-mailout areas.
7. Similarly, contiguous non-mailout areas must be above a specific size threshold and must not contain any small pockets of mailout.
8. Mailout areas close to the minimum size threshold must not be too distant from other mailout areas.

Lastly, the Census project wishes to leverage AR data and the SDI to make field verification as efficient as possible. In 2001, the AR was used to improve coverage in selected areas of the country, and, conversely, 2001 Census data was used to improve the coverage of the AR. In the former operation, results showed that the AR coverage of the chosen areas was very high. From this, the conclusion follows that by choosing a well covered and well geocoded area for mailout in 2006, field updates to the list can be minimized. Based on this conclusion, the following characteristics were also required of the frame:

9. The mailout area should be well covered by the Address Register. Areas poorly covered or not covered will be excluded from the mailout.
10. Using the Spatial Data Infrastructure, addresses must have a high rate of geocoding. If addresses cannot be assigned to blocks, it is not certain those blocks are mailable.

The methodology of delineating the spatial extent of the mailout area is discussed in more detail later.

3.1.2 Required characteristics of questionnaire delivery

Census requirements for questionnaire delivery are slightly different from Canada Post's usual customer requirements. In particular, when people move, they typically have their mail forwarded to their new address. For Census purposes, questionnaires must not be forwarded, as this would destroy the link between address and geography. Canada Post delivery needs to be completely analogous to having an enumerator hand deliver a questionnaire: the questionnaire must be delivered to the address printed on the front of the questionnaire, even if the former occupants have asked for mail forwarding. Undeliverable questionnaires need to be returned to Statistics Canada to allow for more efficient resolution of these dwellings during non-response follow-up.

3.2 Field verification

After the country had been partitioned into mailout and nonmailout areas, the Census needed to field verify the mailing list within the mailout area. For the 2006 Census, this field verification process was called Block Canvass,

and took place between late August and early November of 2005. The mailout consisted of 31,219 collection units, and for each collection unit a map was printed that identified which blocks were in the collection unit, and a corresponding Block Canvass register was printed that contained all addresses from the AR in the collection unit along with the block in which they were located. A total of approximately 9.4 million unique addresses were covered in the 2006 Census Block Canvass area.

The maps contained all information needed for field staff to orient themselves, determine the extent of the collection unit, and update the address frame. All roads were printed, named, and listed with address ranges. The boundaries of legally defined entities, such as municipalities, were printed and the names of the adjacent areas were printed on the boundary. Water features were printed and named.

In addition, blocks within each collection unit were numbered and given a start point to provide an optimal route for the block canvasser through the collection unit. The addresses within the register were printed in the order of the optimal route.

3.2.1 Field operations

The two month Block Canvass operation allowed each block canvasser to canvass many collection units. Thus, for the actual canvassing of 31,219 collection units, Statistics Canada hired a total of about 1900 canvassers. Field staff were assigned zones, which were groups of collection units. When a zone was completed, another was assigned. Within each collection unit, field staff would walk a route indicated on the map, and update the block canvass register, adding missed addresses, deleting addresses that did not correspond to private dwellings within the collection unit, and making minor corrections to existing addresses. In addition, if the map they were provided did not reflect the situation in the field, they also updated the map.

In scheduling Block Canvass a few months before Census Day, Statistics Canada was able to greatly reduce the maximum field staff required at any one time. An additional benefit is that field staff in the mailout area were highly specialized. During block canvass, all they did was list dwellings and update collection maps. During non-response follow-up, all they did was non-response follow-up. In 2001 by comparison, field staff were required to become experts at many tasks, instead of just one. Increased specialization led to better quality in each of these tasks.

3.2.2 Preventing undercoverage

There was some concern that creating the dwelling frame six months before Census Day would increase undercoverage to unacceptable levels. This issue was recognized in advance, and methods were put in place to mitigate this possibility. First, staff were instructed to list all dwellings under construction, so long as an address could be obtained. If these dwellings were not complete on Census day, they would be removed as invalid during non-response follow-up. Second, the front cover of the block canvass register asked for additional information about the collection unit. One box was to indicate the number of dwellings under construction for which no address could be obtained. Finally, staff were instructed to look for evidence of future development, such as signs that a new subdivision or apartment tower would soon be built, and record available information on the front page of the register.

A second field operation, called Late Block Canvass, was conducted in February and March of 2006. Information from the front cover of the Fall 2005 block canvass registers were used to select collection units for this operation. Only dwelling additions could be performed during this operation since the results of the block canvass had already been provided to the printers for addressing the questionnaires themselves. Approximately ten percent of the collection units were recanvassed in Late Block Canvass to ensure the dwelling frame was up to date.

Another risk to the Block Canvass operation was that field staff might complete their registers without actually verifying the information in the field. For the most part, the address lists pre-printed in the Block Canvass registers would be very good. After completing several collection units, field staff might realize that most collection units have minimal change, and some might be tempted to complete the block canvass registers without visiting the

collection unit or while quickly driving through it. To minimize this risk, some addresses were deliberately suppressed from each register and printed on a separate list. Field staff were told that each register was deliberately incomplete, and this approach both reduces the incidence of deliberately poor quality work and increases the ability to detect it. Quality Control technicians ensured that all suppressed addresses were either added back in or correctly left out, and would be able to identify staff who consistently missed the suppressed dwellings.

3.3 Delineating the mailout area

As Canada Post offers mail delivery service to all dwellings in Canada, we would ideally like to mail to all dwellings. In addition to workforce savings already mentioned, another advantage is that it is simpler to conduct a survey with only one collection methodology instead of several. As has already been noted, however, while Canada Post can mail to any dwelling in the country, in more rural areas mail is often delivered by the occupant's name rather than by the address of the dwelling. As the Address Register maintains only addresses and not names, Census mailout can only be conducted where Canada Post uses addresses for delivery.

Second, the mailing list cannot simply be a list of mailable addresses, rather it needs to be the set of all addresses within large contiguous areas made up of blocks. Thus, the Census needed to create a process that built up a mailout area based on the mailability of existing Address Register addresses.

To delineate the 2006 mailout area, the Address Register created from the 2001 Census was used. This version allowed the Census to measure coverage against the 2001 dwelling dissemination counts as each 2001 address also had block information attached to it to ensure accurate geocoding results. Two assumptions were required. First, if it was determined that an area was mailable in 2001, then it would continue to be mailable in 2006. Second, if growth occurred in areas that were considered mailable, then these new addresses would also be mailable. These assumptions were confirmed by Canada Post.

As the mailout area needs to be a complete set of blocks, the first attempt at defining a mailout area was done using 2001 blocks. By determining whether the 2001 addresses were mailable or not, all 2001 blocks could be included in or excluded from the mailout area. This approach did not work well and was rejected for two reasons. The first was that too many blocks have no private dwellings: they are shopping malls, offices, factories, city parks, etc., and thus no evidence exists as to whether these blocks should be included in mailout for 2006. Secondly was that of the remainder, many have low dwelling counts. It is difficult to determine if a block with two dwellings and one AR address should be considered mailable.

Subsequently a mailout area was created using the 2001 Dissemination Area. A Dissemination Area (DA) is another unit containing a small number of blocks. There were 52,993 DAs in 2001. The Dissemination Area proved quite good for delineating mailout, and thus after some experimentation, all DAs for which the AR had better than 95% coverage and for which better than 95% of dwellings were mailable were included in the mailout delineation. This created a largely contiguous set of areas that covered major and minor urban centers, extending to the urban fringe and, depending on Canada Post delivery methods, into rural areas.

While good, this set did not meet all criteria listed in section 3.1.1, in particular the contiguity and size threshold requirements noted in criteria 6 and 7. Thus, spatial algorithms were developed to smooth out the data. In particular, small gaps in the mailout area were filled in and small pockets of mailout were removed. The inclusion of small gaps rests on the assumption that if a small area is surrounded by mailout, then the small area is in fact mailable. As a last step, a manual review was done of all mailout and nonmailout clusters, focussing more particularly on the smaller ones. A small number of areas were altered from one method to the other, mostly nonmailout pockets on shorelines that were converted to mailout.

At this point a mailout area had been created that met all criteria in section 3.1.1. With the spatial smoothing routines, some mailable areas were removed from the mailout, and some areas with poor coverage or data quality were included. As this latter category consisted of small areas entirely within demonstrably mailable areas, the Census was confident of this choice. If this confidence was misplaced, these pockets were relatively small and

could have been handled during non-response follow-up. Post-censal evaluations showed that these pockets of poor quality were mailable.

3.4 Results from the mailout

The 2006 Census successfully mailed out to approximately 9.4 million dwellings. As one Census requirement was that the questionnaire be delivered to the address listed, Canada Post was asked not to forward mail for people who had recently moved. In those cases, questionnaires were either left at the address for new occupants, or the questionnaire was returned as undeliverable, as were questionnaires corresponding to dwellings still under construction.

A number of post office returns were received, but they corresponded mostly to vacant dwellings or to nonexistent dwellings. These cases were resolved by non-response follow-up. A point to emphasize is that there were no concentrated areas of post office returns corresponding to real dwellings that are not mailable. There were isolated cases, but no clusters of dwellings. Thus, the methodology for mailout delineation was successful.

Furthermore, post-censal mailout analysis showed that for many CUs, there was almost no change. For these CUs, the initial frame had a complete set of addresses, with minimal or nonexistent undercoverage, and minimal or nonexistent overcoverage. On the whole, Block Canvass did not significantly alter the number addresses on the frame, nor did it offer significant updates to block information for addresses – most were unchanged. The point here is that the data quality and coverage is not uniform – most CUs were perfect or near perfect, and a small number required significant updating to the frame. A key objective for 2011 then, was to research and implement processes that will further improve the cost effectiveness of field listing, without unacceptable impacts on dwelling coverage.

4. Plans for the 2011 Census

4.1 Goals

Statistics Canada has a number of goals for the 2011 Census related to the mailing of questionnaires. The primary goals for 2011 are to expand the mailout area and to conduct field operations for frame updates only where necessary, a process called targeted listing. Both of these are specific instances of a more general goal, which is to make the Census more efficient and cost effective. Among the minor changes for 2011, the Census is including more areas known to be seasonally occupied dwellings (i.e. cottages) in the mailout area than was the case in 2006.

4.2 Expanding the mailout Area - Working with Canada Post

The goal for 2011 is to mail-out to 80-85% of the dwellings, up from 70% in 2006. The methodology for delineating the mailout area for 2011 is essentially the same as in 2006. Expansion of the mailout area is largely being driven by improvements in address coverage and quality and in improvements in the ability to geocode addresses. Canada Post is also driven to convert additional areas to civic-style address delivery for reasons of efficiency. Statistics Canada is working with Canada Post to choose “conversion” areas, where addressing already exists and could be integrated into Canada Post operations, and where the area would increase the Statistics Canada mailout area. One such site was chosen in Quebec for the 2009 Census test, the results of which will be reviewed by both Statistics Canada and Canada Post to confirm or refine the conversion process. It is intended that this direction be followed in the future, with each successive census seeing increases in the proportion of dwellings covered by the Mailout area.

4.3 Targeted listing

As was discovered in 2006 post-censal analyses, the AR had excellent or perfect coverage of a majority of collection units. Thus, listing in these stable areas is not a productive use of resources. For 2011, the Census will

conduct targeted listing – areas will be classified as being either well covered or not well covered by the Address Register, and listing will be conducted only in non-well-covered areas.

It is not straightforward to identify non-well-covered areas. However, the AR does have access to a number of data sources that help determine which areas are stable and well covered, as well as which are undergoing growth or change. Among these sources are road network updates (as new roads may indicate new dwellings), information from the Canada Mortgage and Housing Corporation which indicate where developers are building, federal tax rebates issued under a New Housing Rebate Program, and electronic phone books. From these sources, the AR can both obtain new addresses and determine spatially where growth is occuring. Both address growth and spatial growth indicators can trigger targeted listing for the affected areas.

Current planning assumptions for 2011 are that Targeted Listing will occur in Collection Units covering approximately 30% of the dwellings in the Mailout area. The 2009 Census Test will evaluate the robustness of the method for identifying areas for Targeted Listing, and will also produce estimates of the anticipated net undercoverage resulting from Targeted Listing in a variety of scenarios. These data will be used to confirm, or revise the targeting method, and also budget and coverage assumptions for 2011.

5. Conclusions

5.1 Conclusions

The Address Register was used in the 1991, 1996 and 2001 Census as a coverage improvement tool. During these coverage improvement exercises, it was noted that the Address Register had very high coverage of urban areas. In 1996, a special project known as the 1996 Centralized Edit Test also showed that the Address Register could be used as a frame for mailing questionnaires, with a field verification operation prior to mailout.

It was also recognized that using Canada Post to mail questionnaires would reduce the field workload required at the time of the Census. This benefit is crucial, as during the 2006 Census, tight labour conditions made it extremely difficult to conduct operations in parts of the country. Had the Census needed to hire the same number of field enumerators in 2006 as in 2001, these problems would have been many times more severe. By engaging Canada Post to deliver 70% of the questionnaires, and by spreading out the listing work for several months, Statistics Canada greatly reduced the number of field staff required around Census Day.

There were many challenges in reaching the mail-out goal for 2006, ranging from developing an entirely new process, to printing a custom booklet for each mailout collection unit, to creating the spatial and address data required for these booklets, to delineating a mailout area for the Census. All challenges were successfully met.

For 2006, the Census project had a goal was of delivering a majority of the questionnaires by mail. Developing this process, while preserving all other goals of the Census, posed a number of challenges. Efficiently delineating a mailout area and creating an accurate dwelling frame for the mailout was made possible by using existing data from the Address Register and Spatial Data Infrastructure. The block canvass operation proved successful in ensuring dwelling coverage for the Census, and the Census successfully mailed out to a majority of dwellings.

Evaluation of the 2006 process has shown that there is no requirement for substantial changes to be implemented to the process of delineating a mailout area. The mailout delineation for 2011 will build on the processes developed for 2006. The evaluation also showed most collection units were minimally changed by field operations. For 2011, the Census will use administrative data to identify where the initial frame requires updating, and conduct targeted field listing only in these areas. These assumptions will undergo scrutiny as part of the 2009 Census Test, with the results serving as input for revisions to process, budget, and field staffing required to meet 2011 Census objectives.

Acknowledgement

The author would like to thank the many people who offered input and suggestions for both my presentation and this paper, but would in particular like to thank Joe Kresovic, Kathy McClean, Pam Tallon and Tim Werschler for their input. This paper is better for their comments and suggestions.