Evaluation of the innovations implemented in the 2009 Canadian Census Test

Danielle Lebrasseur, Jean-Pierre Morin, Jean-François Rodrigue
Jennifer Taylor
Statistics Canada, 100 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6

Abstract

The 2009 Canadian Census Test was officially held on May 12th. It is a key element in the planning of the 2011 Census. Among the innovations tested were a wave collection methodology designed to maximize Internet response, a voice message broadcasted to encourage responses, a process to identify unoccupied dwellings early in the census collection process and a computerized system for field activities management. The sample for the Test consisted in a set of 110,000 dwellings selected in two sites of Canada and a set of 25,000 dwellings selected randomly across the country. The Test sample showed a high Internet response, an improved coverage and a streamlined collection process. In this paper, we will describe some of those new collection methods and present details on the results of their evaluation. We will also give an outlook of how this will be used for the 2011 Census.

Key Words: Census, collection, Internet

1. Context and objectives of the Census Test

Two years prior to each census of population, a test is held in order to put into practice the procedures, methodologies and systems that will be used in the next census. Thus, the main general objective of the Census Test is the evaluation of the operational aspects of the next census, with specific attention given to any changes or new aspects that are planned. For the 2011 Census, the test was held on May 12th, 2009. Some analysis of the results, such as the effectiveness of the methods and the quality of the results, is also performed, but due to the limitations of the sample design of the test, described in the next section, caution has to be used in the interpretation of the results.

There are several new aspects that required special attention at the 2009 Census Test. In this paper, we will describe and analyze the results related to the wave collection methodology, described in section 3, and to the use of a voice broadcast message as a method to encourage responses. Other new aspects that required special attention but not covered in this paper are

- the Dwelling Occupancy Verification (DOV) operation, which is a process to identify unoccupied dwellings early in the collection period
- the Field Management System (FMS), which is a system that, among other things, is used by field staff to manage assignments and update the status of the dwellings in the main system
- the enumerator pay system, which is going from a piece rate to an hourly rate.

2. Sample design

The main sample of the 2009 Census Test was a non-probabilistic sample of 329 Collection Units, which are small geographic areas that typically each contain around 300 dwellings. This amounted to a sample of approximately 110,000 dwellings. The sample was selected in two sites, Montreal (Quebec) and Red Deer (Alberta). Selecting the Collection Units in these areas made it possible to have parts of a large city and a smaller city, to have high growth areas, to have both official languages well represented and to have some challenges for the recruitment of staff, an issue encountered in the last census. The Collection Units were also selected in such a way that we would have areas where we use mail delivery of the census questionnaire, including some for the first time, and areas where we don't. We also attempted to have a proportion of the different major collection methodologies that is similar to what is planned for the 2011 Census.

The first collection methodology, referred to as the Mail-Out-Letter group, contained approximately 70,000 dwellings. These dwellings initially received an Internet promotion letter, asking them to complete their Census Test online, using the Secure Access Code provided in the letter. For those that couldn't or didn't want to use Internet for the Census Test, the letter also provided a phone number that could be used to request that a paper questionnaire be sent to the dwelling.

The second collection methodology, referred to as the Mail-Out-Questionnaire group, contained approximately 30,000 dwellings. These dwellings initially received a paper questionnaire. A Secure Access Code was printed on the questionnaire in case the respondent wanted to complete their Census Test online.

Finally, the third collection methodology, referred to as the List-Leave group, contained approximately 10,000 dwellings. A paper questionnaire was dropped off at their door by an enumerator. A Secure Access Code was also printed on those questionnaires in case the respondent wanted to complete their Census Test online.

A supplementary probability sample of 25,000 dwellings was also selected across Canada. The goal of this extra sample was to do an in-depth analysis of the wave collection methodology and compare different versions of the Internet promotion letter. This supplementary sample will not be part of the scope of this paper. It was covered by Taylor (2009).

3. Wave collection methodology

The wave collection methodology is a streamlined approach of communication with the respondents. It is designed to remind Canadians to complete their questionnaire at key moments of the collection period. It is also designed to maximise Internet response without increasing total non-response. The main reason to strive for Internet responses is that they usually are of better quality than paper responses. Indeed, since there are edits built into the Internet application, inconsistent or erroneous entries are immediately brought to the attention of the respondent for correction, and other inconsistent entries due to skip patterns are also avoided. Another advantage of Internet responses is that they are registered as received almost instantly, as opposed to the few days required for a paper questionnaire to go through the mail process and be received at the processing

center. This contributes to a more up-to-date picture of the collection process status, which in turn provides the chance to allocate resources in a more optimal way.

Figure 1 gives an overview of the wave collection methodology that was used for the 2009 Census Test. The 3 groups of dwellings were subject to up to 4 waves. Wave 1 was a week before Census Test day, wave 2 was 2 days after Census Test day, wave 3 was 10 days after Census Test day and Wave 4 was 20 days after Census Test day.

The Mail-Out-Letter group received, at wave 1, the Internet promotion letter asking them to complete their Census Test online with the Secure Access Code or call to request a paper questionnaire. At wave 2, a reminder letter was sent to the dwellings that had not yet responded. This letter also included the Secure Access Code and the phone number to call to request a paper questionnaire. The dwellings that had still not responded by wave 3 then received a paper questionnaire with a letter asking them to complete the Census Test before May 31st using the enclosed questionnaire or go online and use the Secure Access Code provided. This letter also mentioned that if the questionnaire was not completed and returned by May 31st, a census enumerator may contact the household by phone or in person to help complete one. This is leading us to wave 4, which is the Non-Response Follow-Up (NRFU), held starting 20 days after Census Test day until mid-July. As explained in the wave 3 letter, this is when census enumerators attempt to contact the non-responding households by phone or in person to complete questionnaires.

The Mail-Out-Questionnaire group received a paper questionnaire at wave 1, and a Secure Access Code was printed on the questionnaire in case the respondent wanted to complete their Census Test online. The other waves were almost identical to those of the Mail-Out-Letter methodology, with one difference at wave 3. Instead of all the non-respondent dwellings at that point receiving the letter and paper questionnaire, the group was split into two sub-groups. A pre-determined half of the 30,000 dwellings would receive the letter and a second paper questionnaire while the other half would receive a voice broadcast reminder message over the phone. The results of the comparison of these two follow-up methods are analyzed in section 5 of this paper.

Finally, for the List-Leave group, wave 1 consisted of enumerators dropping off a questionnaire at the door of the dwellings. A Secure Access Code was also printed on those questionnaires. At wave 2, an unaddressed reminder ad-card was distributed to these dwellings and NRFU started at wave 3.

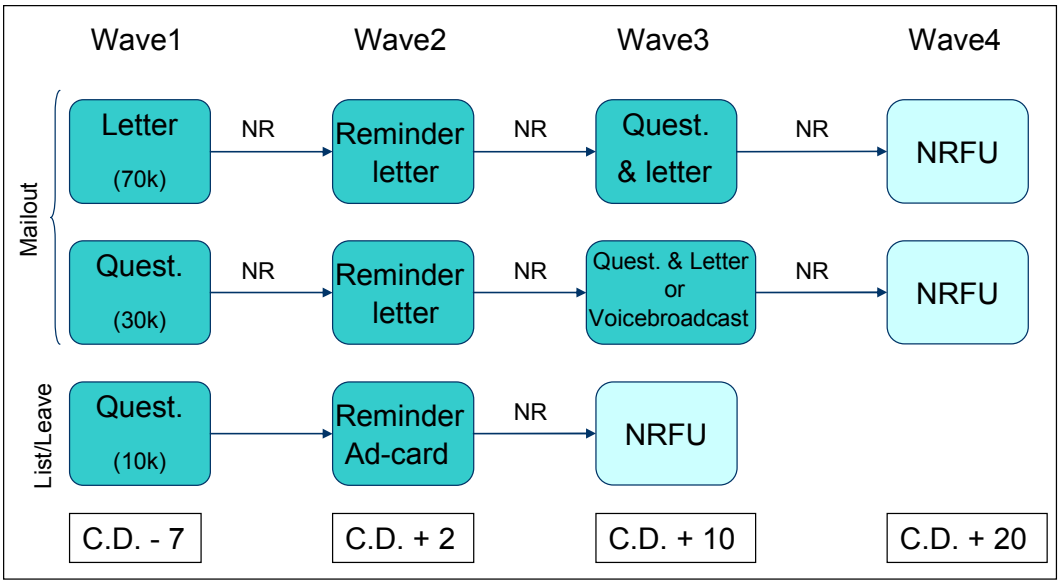

Figure 1: The wave collection methodology for the 2009 Census Test

4. Results from the main sample

The overall response rate for the Census Test's main sample was 77.3%. Another 0.5 percentage point was obtained through the Census Help Line, but these responses are not taken into account in this paper. As shown in Figure 2, the response rate was almost identical across all collection methodologies. However, there is a large difference in the distribution of the response channel (mail, Internet, field) between the Mail-Out-Letter methodology and the other two methodologies.

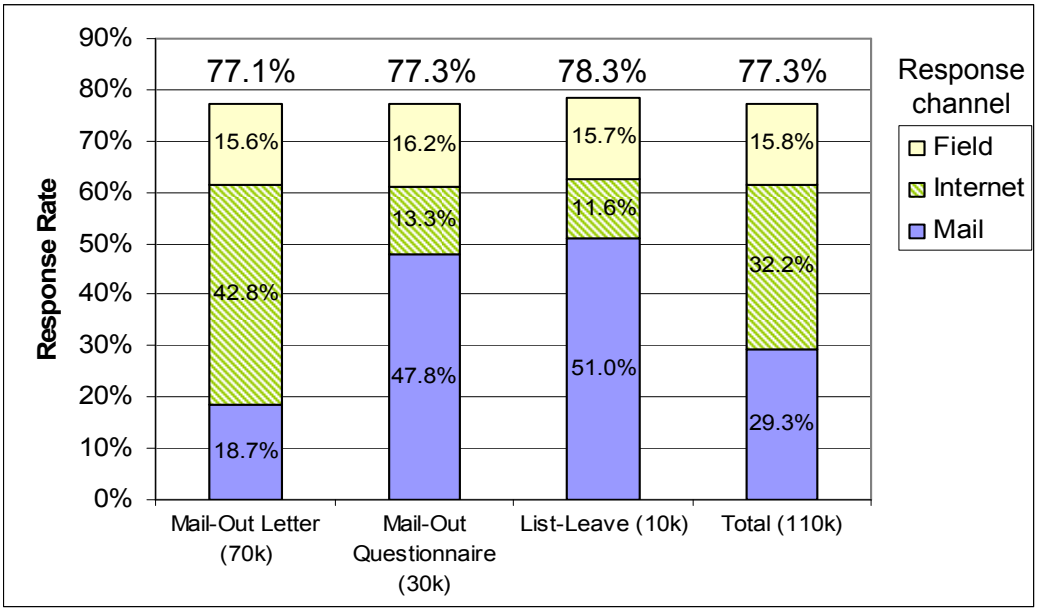

Figure 2: Response rates by collection methodology

Indeed, 42.8% of the dwellings in the Mail-Out-Letter group responded via Internet, representing 55.5% of the respondents of that group, while only 18.7% responded by

mail. Conversely, in the Mail-Out-Questionnaire and List-Leave groups, only 13.3% and 11.6% respectively of the dwellings responded by Internet, but 47.8% and 51.0% responded by mail. While it is important to mention that the dwellings in the Mail-Out-Letter group were chosen because they are in areas that had high response rates in the last census and with a high potential of Internet users, the wave collection methodology clearly has an impact on the response channel used by the respondents. For comparison purposes, at the 2006 Census there were 18.3% of the respondents that completed their census online, and that was the highest rate ever achieved for a census in any country up to then (Laroche, 2007). It is however difficult to tell, with the results of this non-probabilistic sample, how much the wave collection methodology affected the overall response rate.

Figure 3 shows the evolution of the responses during the collection period. Each curve represents a response channel, and therefore the total of the three curves at any point amounts to the total response rate for the Census Test at that point. We can see that the Internet response started very strong, and then stayed on top until the end. The strong start for Internet responses can be partly explained by the fact that Internet responses are registered almost instantly when the questionnaire is completed. On the other hand, mail responses came in at a steadier pace. And of course, we didn't get any significant field response until the start of the NRFU operation for the two Mail-Out groups, about three weeks after Census Test day. The columns for the different waves are located where we should start seeing a significant effect of the wave on the response. We can see some sort of plateau just before wave 2 and wave 3, especially for mail responses. Also, mail and Internet responses are still being received after the start of the NRFU operation. That gives us an indication that waves 2 and 3 may be a few days too late and we should look into adjusting the wave calendar for the 2011 Census. But no major operational issue was found.

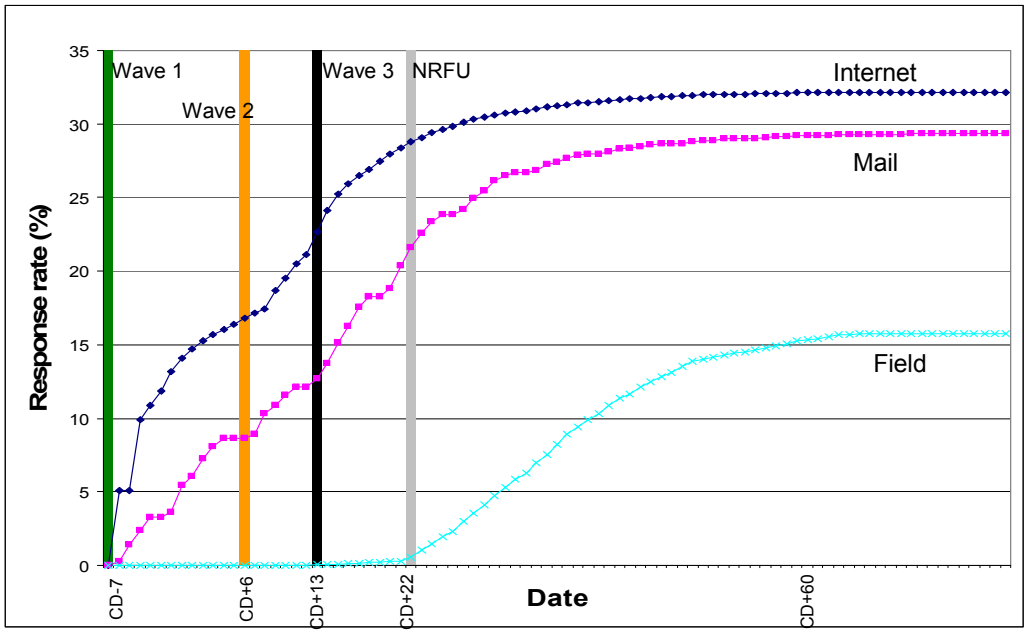

Figure 3: Cumulative responses by response channel

5. Two different follow-up methods

As part of the objectives for the 2009 Census Test, we were also looking for an alternative to sending a second questionnaire to non-respondent dwellings at wave 3. Hence, a study was done by randomly splitting the Mail-Out-Questionnaire group into two sub-groups. A pre-determined half of the 30,000 dwellings would receive a second paper questionnaire if they were a non-respondent after wave 2 while the other half would receive a voice broadcast reminder message over the phone.

The phone numbers used for the voice broadcast message came from Statistics Canada's Address Register, which is the frame for the Mail-Out component of the census. They were extracted from administrative sources such as telephone billing files and were available for about half of the dwellings.

Both methods have advantages and disadvantages. Sending a second questionnaire is a widely used method and is known as being effective. However, it is expensive and logistically complicated. Indeed, the questionnaires have to be addressed and mailed over several days. On the other hand, a voice broadcast message is an almost instant process and is relatively inexpensive. However, expectations for it were low since in general automated voice messages received at home do not get a lot of attention and are not popular.

Figure 4 compares the response rates for the dwellings that had not responded by wave 3, by voice broadcast status. This does not include responses obtained through NRFU. For comparison purposes, this figure also shows the response rate of the dwellings that received a second questionnaire.

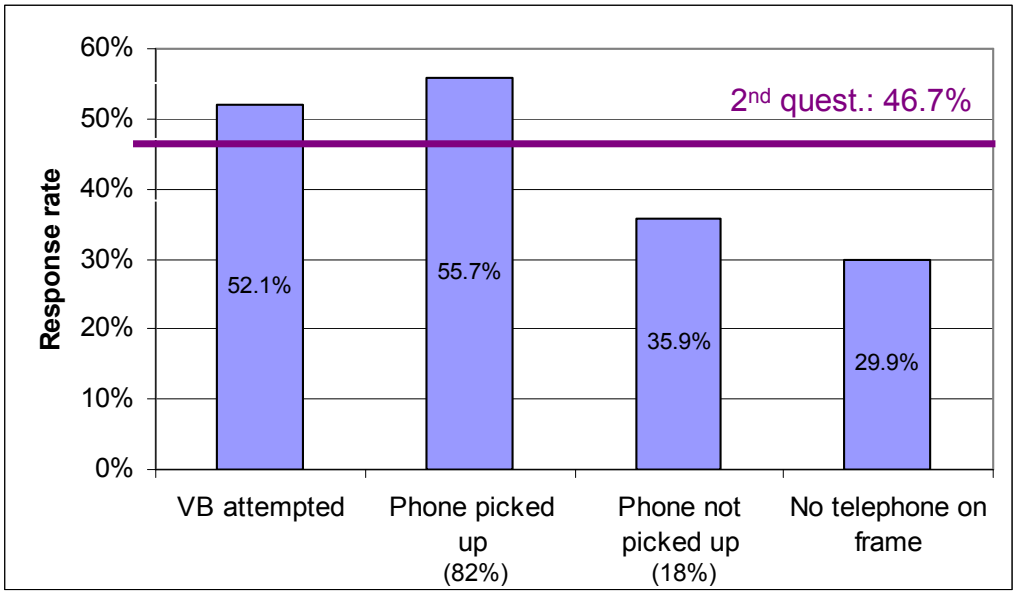

Figure 4: Response rates by voice broadcast status

The first column shows the response rate for the dwellings where the voice broadcast was attempted. The next two columns show the breakdown of the result of that first column in whether the phone was picked up (including answering machine and voicemail) or not. It is worth mentioning that when the voice broadcast was attempted, the phone was picked up 82% of the time. Finally, the last column shows the result for the dwellings where we

should have sent the voice broadcast but for which no phone number was available on our frame. The response rate for the dwellings where we sent a second questionnaire was 46.7%, which is surprisingly lower than that of the dwellings where the phone was picked up for the voice broadcast (55.7%) and even lower than that of all dwellings where the voice broadcast was attempted (52.1%).

These unexpected results led us to wonder if our two groups were comparable. One known difference is that the dwellings where we could attempt a voice broadcast were only the dwellings for which we had a phone number on our frame, whereas we were able to send a second questionnaire to all of the dwellings from the other group. Therefore, Figure 5 splits the dwellings that received a second questionnaire into whether or not we have a phone number for them on our frame. It is interesting to see the large difference between them.

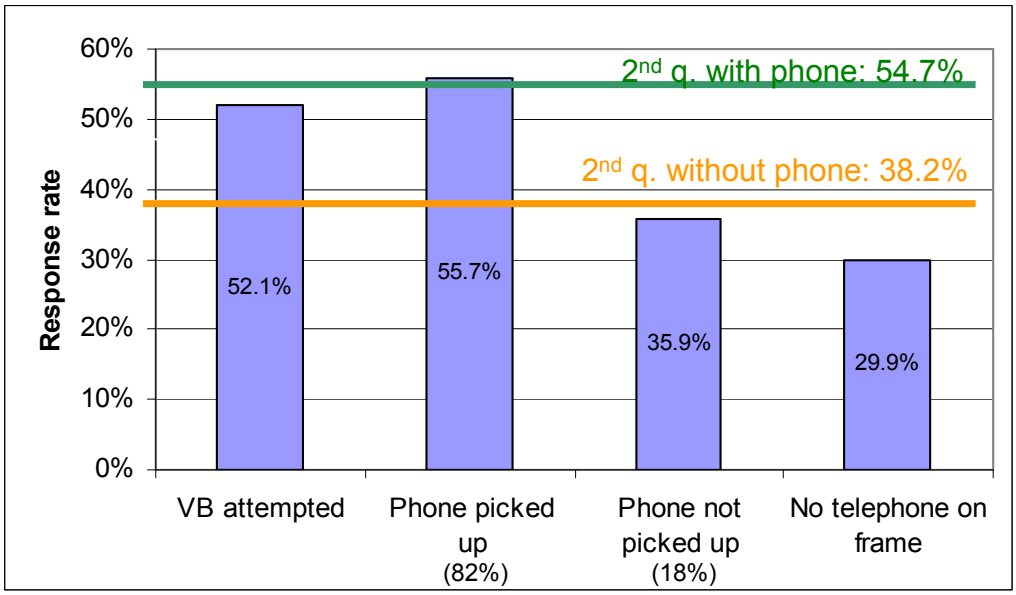

Figure 5: Response rates by voice broadcast status, with second questionnaire group split

Indeed, the response rate was over 16 percentage points higher for the ones for which we have a phone number on our frame. The response rate for that group was actually approximately the same as what was obtained for the dwellings where the phone was picked up for the voice broadcast and slightly higher than what was obtained for the dwellings where the voice broadcast was attempted. Therefore, it seems like the presence of a phone number on our frame is a relatively strong indicator of the probability to respond. This could also explain the difference between the response rate of the dwellings where the voice broadcast was attempted but the phone was not picked up and that of the dwellings where the voice broadcast was not attempted due to the absence of a phone number on our frame.

Despite the fact that the dwellings that received a second questionnaire and for which we have a phone number on our frame did slightly better than the ones where we attempted the voice broadcast, we can certainly say that the voice broadcast method exceeded expectations. Furthermore, a follow-up survey was carried out on a sample of the dwellings that received the voice broadcast and, among those who responded and remembered receiving it, only 2% said that it had a negative influence on their

participation in the Census Test. Also, the information in the message was considered not sufficient by only 2% of the dwellings, so we can say that the voice broadcast was efficient at stimulating more responses. However, it is worth noting that the sample of this follow-up survey might be biased towards positive results since respondents answering such a survey might be more inclined to have a positive outlook on surveys in general. Therefore, it is recommended to interpret these results with caution.

6. Summary

The wave collection methodology seems to have worked well. No major issue was found, operationally or in terms of response, and we saw promising results for Internet response. However, an in-depth analysis of the supplementary sample is necessary to be able to precisely assess the impact of the methodology on the NRFU workload and to be able to introduce refinements to the strategy.

Also, the test demonstrated that a voice broadcast message is an efficient and relatively inexpensive follow-up method that can be used for dwellings for which we have a phone number on our frame. Furthermore, we realized that the presence of a phone number on our frame is linked to characteristics of dwellings that respond more. That is a powerful indicator to have on a frame.

7. Towards the 2011 Census

Based on the above conclusions, a preliminary plan for the 2011 Census collection methodology has been put together. The preliminary plan will use the wave collection methodology and the Internet promotion letter similarly as what was done for the 2009 Census Test, but with the modifications specified in the non-colored boxes in Figure 6. More precisely, we will modify the wave calendar by moving wave 2 and wave 3 earlier by two days. Also, we will use a voice broadcast message at wave 3 for the Mail-Out-Questionnaire group. Finally, a voice broadcast message will also be sent just before the beginning of NRFU and also during NRFU.

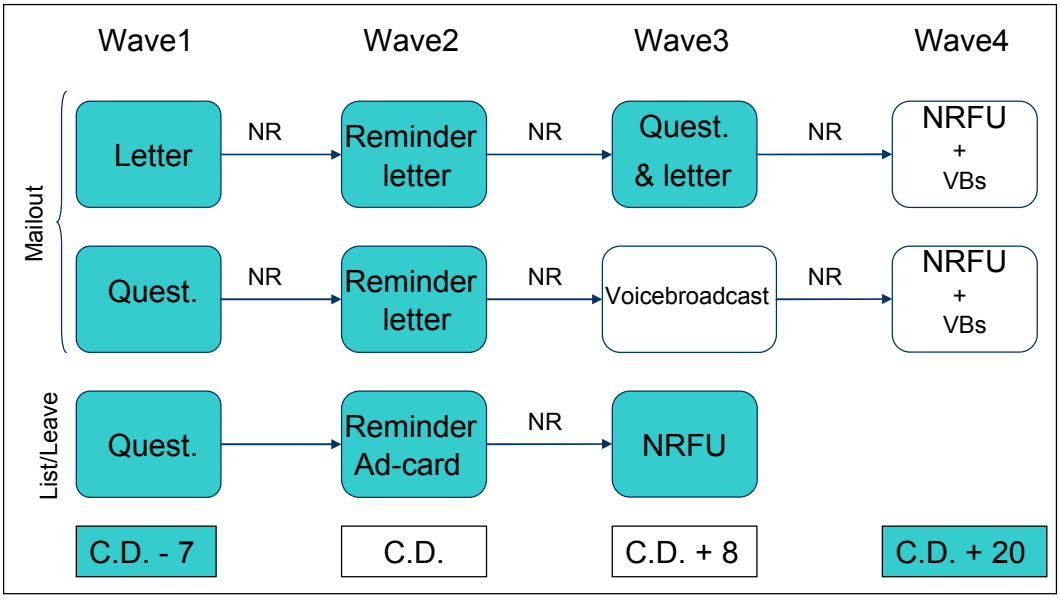

Figure 6: The planned wave collection methodology for the 2011 Census

Acknowledgements

The authors would like to thank Patrice Mathieu and Laurent Roy for their helpful comments.

References

Taylor, J. (2009), "2009 Census Test: Analysis of the supplementary sample", Internal document, Statistics Canada.

Laroche, D. (2007), "2006 Census Internet data collection", Advisory committee on statistical methods, Statistics Canada.