



Economic and Social Council

Distr.: General
13 March 2012

Original: English

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Population and Housing Censuses

Fourteenth Meeting

Geneva, 24-25 May 2012

Item 5 of the provisional agenda

Internet data collection

Internet data collection in the Canadian census of population

Note by Statistics Canada

Summary

The 2011 census of Canada saw the introduction of several methodological changes. While building on the successes of the 2006 census when were introduced central mail-out with the use of an address register and an internet response option, many of the changes in 2011 were made to promote a higher rate of return via the internet while reducing risk of higher non-response. The main element of the approach was a wave methodology. The methodology consisted of repeated mail, telephone and face to face reminders to non-respondents to encourage them to primarily self-enumerate, and to do it by internet as a preferred mode. This method was tested successfully in 2009 before introduction in 2011.

I. The Canadian Census

1. This document was prepared by Marc Hamel, Statistics Canada.
2. Canada conducts a census of population and dwellings every 5 years. The most recent one took place in May 2011 and saw the introduction of several methodological changes. While building on the successes of the 2006 census when were introduced central mail-out with the use of an address register and an internet response option, many of the changes in 2011 were made to promote a higher rate of return via the internet while reducing risk of higher non-response. The main element of the approach was a wave methodology. The methodology consisted of repeated mail, telephone and face to face reminders to non-respondents to encourage them to primarily self-enumerate, and to do it

by internet as a preferred mode. This method was tested successfully in 2009 before introduction in 2011.

II. The census collection methodologies

3. The Canadian census uses different methodologies in conducting its collection. They are briefly described in Rodrigue et al (2012). The main methodology consists of a mailout using the information available in an Address Register. This methodology was applied to approximately 79 per cent of all dwellings in 2011. For the others, a questionnaire was either delivered to the dwellings by a census enumerator as part of a list and leave operation (19 per cent), or the census was collected directly at the door by a census enumerator (2 per cent). Collection at the door is mostly used for First Nations communities, and those in remote and northern areas. Central mailout is not used more widely because the national postal service cannot deliver unaddressed mail to all dwellings. Statistics Canada does not have names or residents associated with dwellings as part of the Address Register. Since the dwelling is the basic starting point for the census enumeration of the population in Canada, it is imperative to provide specific letters or questionnaires to specific dwellings to determine who has responded ensure complete enumeration. The list and leave approach is used mostly in rural areas where mail is generally not delivered at the door but to a post-office box and requires a name to be delivered to the right dwelling.

III. The wave approach

4. The wave approach used for the 2011 Canadian census is largely based on the theory of Dilman (2007). The approach was applied differently based on the main collection methodology. There were two different approaches for mailout areas, and one approach for list and leave areas. The choice of approaches in mailout areas was determined based on the propensity of the population to self-enumerate as demonstrated in the 2006 census and the 2006 Census internet take-up rate.

IV. Approaches in mailout areas

5. For approximately 75 per cent of dwellings in mailout areas, a letter only was sent in wave 1. This treatment was identified as group 1. Letters were all delivered on 3 May by the postal service. It invited respondent to respond online using the personal Secure Access Code printed on the letter, and also included a toll free telephone number to order a paper questionnaire if this was preferred (calls were made to the Questionnaire Request System). The toll free number of the Census Help Line was also included in case respondents had any questions or required assistance. This approach which did not offer a simple option to answer on paper was used in areas where self-enumeration was likely to be higher. The remaining 25 per cent of dwellings in the mailout universe were mailed a questionnaire package. The package also included a personal Secure Access Code for the online option and the number to the Census Help Line. This approach was identified as group 2.

6. Wave 2 started on 10 May with the production and mailout of a reminder letter to all non-respondents from wave 1. The letter was in the same format as the letter from wave 1 and indicated to respondents that it was not too late to respond. It also included the dwelling's Secure Access Code again so that the initial letter or questionnaire package would not be required to respond online. Both groups received the same letter for wave 2. Letters were printed using an 'on demand' process in a self-mailer format where the letter

folds in what is also the envelope (one piece of paper). This format allowed for the production of more than 6 million reminders over a 6 day period. The process also allowed for the matching of the address and the Secure Access Code for each letter, critical to ensure that a return is accounted for the right dwelling.

7. Wave 3 started on 18 May, 8 days after census day. For group 1, a questionnaire package was sent to all dwellings for which a response had not been received and a questionnaire had not been ordered via the Questionnaire Request System. This package contained a new access code for internet. The letter accompanying the package contained more direct wording concerning the mandatory requirement to complete the census. Again a matching process was put in place to ensure that the address and the Secure Access Code were associated to a specific dwelling.

8. For the group 2 that had already received a questionnaire package in wave 1, a phone message was sent to non-responding dwellings using a voice broadcast service. The message was a reminder to complete the census either online or by mail and was sent in mid-day in an attempt to be left on an answering machine. Telephone numbers were available for approximately 60 per cent of dwellings, mostly from publically available sources.

9. Wave 4 started on 1 June and consisted of telephone or personal follow-ups of remaining non-respondents by enumerators. If personal contact was not established during initial follow-up, either a message would be left on the answering machine or a 'Notice of Visit' card would be left at the door indicating the purpose of the follow-up and asking the householders to call back the Census Help Line or the nearest Local Census Office for assistance in completing the census. Follow-ups continued until response rate objectives were met for all areas.

V. Treatment in list and leave areas

10. In list and leave areas, questionnaire packages were delivered to all dwellings by Statistics Canada enumerators in wave 1. Delivery took place from 2 May to 10 May. The packages included a questionnaire and a personal Secure Access Code to the online application.

11. Wave 2 consisted of a general reminder card to all dwellings as part of an ad-mail delivery. There was no means of sending the reminder only to non-respondent dwellings.

12. As there was no treatment planned as wave 3 for these areas, non-response follow-up started on 20 May, 10 days after census day. The early start to non-response follow-up was consistent with the approach used in the 2006 Census and ensured that the enumerators used for the drop-off of questionnaires would mostly still be available for follow-up.

VI. Questionnaire Request System

13. The Questionnaire Request System is an automated telephone system that was put in place to allow households who had received only a letter in wave 1 or 2 to order a paper questionnaire if this was their preference. After calling the number available on all letters, respondents would be prompted to enter their Secure Access Code using their touch tone phone. The system would then send a message to Canada Post to address a questionnaire package to be mailed to the respondent. The average return time for completed paper questionnaires issued in this manner was 12 days. Respondents without a touch tone telephone were routed to the Census Help Line for assistance.

VII. Testing the wave approach

14. The approaches used in 2011 were based on the results of tests conducted in 2009. The census test conducted on a sample of 100,000 dwellings used the basic approach (4 waves) and demonstrated that it could be used successfully (Taylor 2010). In fact the internet response for the test which was collected on a voluntary basis was 77.8 per cent, far exceeding the results of the 2004 test. Levels of self-response were also higher than expected at 61.5 per cent (the plan was 47.5 per cent for the test).

15. Four variations of the approach were also tested on samples of 5,000 dwellings each and results were measured against a control group (Taylor 2010). Two different wave combinations were tested, each with two different letters in wave 1. One combination (two separate samples) included a letter in wave 1, a reminder letter in wave 2 and a questionnaire package in wave 3. The second combination replaced waves 2 and 3 with a questionnaire in a wave 2 which took place later than wave 2 in the first combination. No field follow-up was conducted on any of these samples. The test demonstrated that the approach involving 3 mailouts as in the first combination was the most promising to generate high internet response while minimising risks of increased non-response and was selected for 2011.

16. This test also measured results when using two different wordings for the wave 1 letter. One letter was inviting respondents to respond on census day (10 May) while the other was inviting them to answer within 10 days of receipt. The two letters were used in each combination, one for each sample. The test did not show a material difference related to wording. Based on this, the 10 days message was selected in the hopes of spreading the internet response over several days and avoid too high of a peak on the infrastructure on a single day. This would reduce the risk of infrastructure failure. Since Canada uses a *de jure* approach for its census, it was not felt that inviting respondents to respond within a 10 period would cause issues of precision on the population count. In past censuses, questionnaires were always delivered in advance of census day and a sizeable portion of households would complete them before census day.

VIII. Response management

17. A strategy was put in place to manage the collection process to ensure that the target response rate was going to be reached, but also to ensure as uniform response levels by region as possible. It was put in place to make optimum use of the human and financial resources in the pursuit of response objectives. The strategy required the availability of management information on the progress of collection almost in real time. It also made use of a dynamic model to indicate the projected end of collection by region based on a number of parameters – self-response levels, productivity of enumerators in non-response follow-up, and the hours worked per day by enumerator by area. This information was analysed several times a week during collection to determine if different tiers of the collection management strategy were going to be invoked.

18. The strategy was comprised of four tiers. It was applied individually to more than 45,000 collection units in the country. These units comprise the collection geography for the planning of collection operations. Tier 1 called for the stop of non-response follow-up in any collection unit across the country if an observed collection response rate of 98 per cent was reached. It was felt that at this level high quality information would be available for dissemination. It did not mean that response would be capped at 98 per cent since self-completed questionnaires would likely keep coming in by internet or by mail for these areas.

19. Tier 2 was to stop collection below 98% in very small collection units (just a few dwellings) as this would have a very marginal impact on the dissemination results. Tier 3 was to stop collection at the response level obtained in 2006, and Tier 4 was to introduce any measures that would ensure highest response possible below 98 per cent or 2006 results. As collection units reached their response target or where collection was stopped, collection resources were re-allocated to neighbouring collection units that were not completed yet. Tiers 2 to 4 were to be invoked only if it was determined that collection in an area would not be completed by the planned end date, or that the budget for collection would all be expended. For 2011, tiers 3 and tier 4 were not invoked.

IX. Response results

20. The various strategies and collection methodologies that were put in place ensured that the response objectives for the 2011 census were reached. The table below shows the national and provincial global collection response rates, the internet collection response rates and the self-response collection rates.

Table 1 (per cent)

	<i>Response</i>	<i>Internet</i>	<i>Self-response</i>
Canada	98.1	54.4	84.1
Newfoundland	98.2	37.6	84.9
Prince-Edward-Island	98.3	36.4	85.8
Nova Scotia	98.0	42.2	85.4
New Brunswick	97.7	48.2	85.3
Quebec	98.2	52.3	84.9
Ontario	98.3	57.0	84.7
Manitoba	97.8	47.4	83.1
Saskatchewan	97.8	41.8	82.7
Alberta	98.0	57.7	82.1
British Columbia	97.6	60.5	83.3
Yukon	94.1	35.5	58.4
Northwest Territories	96.6	9.1	26.2
Nunavut	92.7	0.0	0.4

Notes: These preliminary rates are obtained directly from collection results, i.e. before data processing and data quality verification. They are calculated as the number of private dwellings that returned a questionnaire divided by the number of private dwellings classified as occupied by field staff. After processing and quality verification of the data, post-collection response rates will be produced. Among other improvements, these final response rates will include collective dwellings and adjustments to the number of occupied private dwellings based on a sample study of the quality of the dwelling occupancy status.

21. The wave collection approach was successful in generating not only high response by internet, but also most likely high self-response. The use of a letter only in waves 1 and 2 for a large portion of the census universe pushed respondents to massively use internet as their mode of response. The frequent reminders also seemed to have had the effect of pushing to action people who might have otherwise waited for the visit of the enumerator.

22. Post-collection qualitative evaluation of the respondent messages, including the wave letters and the paid advertising, have shown that Canadians interviewed were generally aware of the census and understood that they had to respond. It is not clear if the controversy around the replacement of the long-form for 2011 Census by a voluntary survey (National Household Survey) and the related media coverage played a role in the level of awareness.

23. When looking at response patterns with the wave approach, we have observed an increase in response level approximately at the time new reminders were sent to non-respondents (Rodrigue et al, 2010, graphic not included). Because the waves may have overlapped more in some regions (fewer days to react to a reminder), it will be difficult to measure the actual impact level of each wave. We can conclude though that sending a letter to a portion of the dwellings had the desired effect (table 2). Group 1 mostly responded by internet (71.6 per cent). For group 2, the mail response level (50.1 per cent) shows that respondents will use different ways of responding when they have readily available options.

24. These results are also logical in the sense that group 1 was comprised of regions with good self-response levels in 2006. Results still exceeded expectations globally as the target global response was 98 per cent, the internet target response was 40 per cent and the target self-response was 80 per cent.

Table 2 (per cent)

<i>Collection method</i>	<i>Response mode</i>					<i>Total</i>
	<i>Mail</i>	<i>Internet</i>	<i>Help line</i>	<i>Non-response follow-up</i>	<i>Non-response</i>	
Group 1 - letter	16.1	71.6	0.7	9.1	2.3	100
Group 2 - questionnaire	50.1	25.8	0.8	20.0	3.4	100

X. The benefits of using Internet as a response mode

25. Response by internet offers a number of benefits. Responses received on the internet tend to be of much better quality as the application guides the respondents through all the relevant questions. Although a respondent does not need to answer any of the questions to move to the next one, item non-response tends to be very low. Table 3 compares the item non-response between the 2006 and 2011 censuses. The reduction of item non-response in 2011 is entirely due to the higher proportion of responses received by internet.

Table 3

<i>2011 question number</i>	<i>Questions</i>	<i>2006 non response rate</i>	<i>2011 non response rate</i>	<i>Difference (2011-2006)</i>
CENSUS				
2	Sex	1.4	0.8	-0.6
3	Date of birth	0.7	0.5	-0.2
4	Marital status	2.9	2.5	-0.4
5	Common Law Status	6.7	4.4	-2.3
6	Relationship to Person 1	1.7	1.4	-0.3
7	Official languages	n/a	1.6	n/a
8A	Home language	n/a	1.8	n/a
8B	Other languages	n/a	2.0	n/a
9	Mother tongue	2.0	1.9	-0.1
	92 Years Consent	8.1	5.5	-2.6

26. Internet also simplifies the processing of returned responses. Data capture is pushed at the respondent level, and the data coming in is already in machine readable format. Because it tends to be cleaner, it also facilitates coding and editing operations when required.

27. With the reduction in the massive use of paper for questionnaires, internet is also a more environmentally friendly option. With the reduction in printing, mailout, mailback, and some of the processing requirements, the internet option can also offer a reduction in costs if the take up is of certain proportion. Canada is currently assessing what this breakpoint is.

XI. Lessons learned

28. The approaches used to increase internet response in the Canadian census have proven to be a success. There were still a number of important lessons learned which will lead to possible improvements for future censuses. The time required to produce and get the reminders to dwellings in various parts of the country took slightly more time than anticipated, which meant that we did not fully benefit from the effect of a wave before initiating the next. Statistics Canada will be assessing ways to implement a tighter application of the approach for the future, looking for ways to produce and mailout each wave in a much shorter time frame.

29. Using this approach also requires that the mailing list be of very high quality. We had a small number of duplicate addresses for some dwellings which actually were not multi-unit structures. In these cases, the respondent generally completed one of the questionnaires and ignored the other, which meant that they kept receiving reminders for the one not completed. These situations were not resolved until wave 4 when a census enumerator was able to verify that the second dwelling did not actually exist. Processes will be put in place to improve the quality of the mailing list, but also to deal more efficiently with such situations during collection.

30. The post census evaluation of the communications material revealed that many Canadian do not make extensive use of traditional mail services, which does not make the wave approach as implemented in 2011 very effective for them. These people manage most of their financial obligations online, and do not have a requirement to often go to their mailboxes. Although the size of this segment of the population is unknown, we can anticipate that it is only going to grow in the future. Statistics Canada will be evaluating options to deliver invitations to participate in the census in different ways for the future, such as inviting respondents to pre-register to complete the census online, or possibly using the e-post service offered by Canada Post to reach some households. With the e-post service, Canadians can register to obtain an electronic address from Canada Post, and have their bills and other similar mail sent to them via that address.

XII. Conclusion

31. The use of the wave approach for the 2011 Census of Population was highly successful. It will be relatively simple to introduce improvements in future censuses to increase Internet response at even higher levels. This will offer opportunities to optimise related processes even more for future censuses such as in data processing and quality assurance.

Bibliography

Dillman, D.A., "Mail and Internet Surveys: The tailored Design Method, second edition", Wiley 2007

Taylor, J., "2009 Census Test: analysis of the supplementary sample", internal document, 2010

Taylor, J., "Étude de la nouvelle method de collecte de recensement dans le cadre du test du recensement de 2009", Recueil du Symposium de 2010 de Statistique Canada, 2010

Rodrigue, J.-F., Mathieu, P., Morin, J.-P., Taylor, J., « La méthodologie de collecte du Recensement et de l'Enquête nationale auprès des ménages de 2011 au Canada », Paper presented at the Journées de méthodologie Statistiques conference, Paris, France, January 24-26, 2012

Côté, A.-M., Laroche, D., Wang, Q., "Identification des unités de collecte qui ont reçu la lettre destinée à promouvoir la réponse par internet au recensement de 2011", Assemblée annuelle de 2011
