United Nations ECE/CES/GE.41/2015/5

Economic and Social Council

Distr.: General
15 July 2015

Original: English

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Population and Housing Censuses

Seventeenth Meeting
Geneva, 30 September to 2 October 2015
Item 3 of the provisional agenda
Innovations planned for 2020 census round, and results of tests

Plans for the 2016 Canadian Census, related to tests

Note by Statistics Canada

Summary

Canada conducts a Census of Population every 5 years. The next Census will be in May 2016. This frequency presents both great opportunities to accelerate the introduction of innovations for all aspects of the program, and challenges in finding sufficient time to conduct elaborated tests. One approach that has been used is to design controlled live tests conducted during census operations, in preparation for the next cycle.

This paper will briefly describe the methodology for the 2016 Census. It will then describe the tests conducted in preparation for the census and some of the findings. It will also describe the live tests planned for production in 2016. Finally, it will briefly describe the work planned to increase the use of information from administrative files on the Census program, and some of the planned research and tests to verify anticipated applications of such sources.

I. Introduction

1. Statistics Canada is responsible under the Statistics Act for conducting the Census of Population every five years. The Canadian Census Program includes the Census of Population questionnaires with approximately 10 questions on basic demographics, and the National Household Survey collecting more detailed socioeconomic information on a sample of households. The program relies on a combined approach, with data collected from individuals and about housing units at a specific point in time and some information being obtained from administrative records (for example: income variables are created using income tax records). Conducting a census every five years provides an opportunity to accelerate the learning on innovative approaches and key changes to any element of the program. It also requires discipline in relation to the timeliness of the necessary tests to support the implementation of such innovations. You basically have to plan larger changes for a future cycle as you conduct the current one.

2. As part of Statistics Canada's customary process, all major changes to methodologies and approaches to important statistical programs are thoroughly tested before implementation. On the Census of Population Program, content and operational tests are usually conducted in the 2 or 3 years preceding the main collection itself. For the 2016 Program, a content test (BT2) was conducted in May/June 2014, and a test of the operational changes (BT3) was conducted in October 2014. Qualitative testing (BT1) of the questionnaire content and instruments was performed before BT2.

3. Particular to the Census Program, new approaches are also sometimes tested in controlled environments during census operations, in preparation for the next cycle. Such tests are now being designed for 2016.

4. This paper presents a brief description of the tests conducted or planned in relation to the 2016 Census Program. Key findings are provided for tests already conducted. Research assumptions are presented for upcoming tests. Statistics Canada is also currently conducting some research to increase the use of administrative data on the Census of Population Program. The paper briefly describes some of the related research projects and planned tests in relation to the 2016 program.

II. Behavior Tests for the 2016 Cycle

5. Over the last few years, Statistics Canada has undertaken a redesign of its global business architecture. The underlying objective of the redesign is to rationalise investments in infrastructure and corporate processes by developing global approaches as opposed to local ones for individual programs. The Census Program is included in these initiatives, especially from the perspective that the large investments required in the program can be leveraged to support the development of corporate initiatives. For example, the field operations systems used for the Census can be expended to meet the requirements of all other survey programs, reducing the required corporate investments in these systems, and also reducing the number of systems and platforms to maintain over time. To make it possible for the Census Program to use corporate tools in 2016, the Census testing strategy was redesigned to focus primarily on changes since the last Census in 2011, as opposed to conducting a full dress rehearsal. The testing strategy was separated in smaller behavioural tests (BTs) to verify the behaviour of respondents and field personnel in relation to new systems and approaches.

6. Questionnaire content and instruments were tested in Behavioral Test 2 (BT2) in May and June 2014. In particular, the new corporate electronic questionnaire instrument was used during the test, replacing the Census specific application used since the 2006 Census. The Census testing capacity was utilised to test the applicability of the electronic questionnaire on different technological platforms, such as desktops, tablets and smart phones. With an objective of 65% for internet uptake for 2016 and rapid changes in technology, we needed to validate that the electronic questionnaire was going to perform adequately on all platforms likely to be used by respondents in 2016. The test confirmed that the objective for internet uptake is appropriate and that the questionnaire performs as designed on all platforms. Another smaller platform test will be performed at the end of 2015 to verify that the electronic questionnaire will work on latest technology.

7. Behavioral Test 3 (BT3) was conducted in the fall of 2014 to test the use of new collection systems and processes by field personnel and the related training material. In particular, we tested the new corporate Integrated Collection and Operation System (ICOS). This system replaces the Field Operations System used in the 2011 Census and provides an internet interface to field personnel to receive work assignments, record work progress and related pay and expense information. The systems also include a number of related functions such as messaging, management information reports, etc. BT3 was separated from BT2 and conducted later contrary to past practices on the Census to allow more time to complete the development of the proper ICOS functionality to be used by the Census.

8. The test of the ICOS showed that certain functions built to be generic (i.e. to be used by regular interviewers on all surveys and by census enumerators) were too complex in a Census context, where it is desirable to limit the amount of training required/provided for cost reasons and because work assignments are limited to short periods of time, requiring a fairly short learning curve. Based on the results of this test, some of the generic functions are being simplified to be used for the 2016 Census. An example of this is the multiple functions developed for field supervisors to create enumerator assignments, such as based on the average number of attempts per case, the proximity of cases to the enumerator's home, etc. These functions proved to be too complex in the context of the Census.

9. A number of adjustments and changes are being made to systems based on the results of the BTs. To validate that these changes will efficiently address the challenges observed in testing, Statistics Canada will be conducting another small systems validation test in the fall of 2015. Lay persons will be hired to conduct this test to verify that the systems will be intuitive in the context of the census operations when more than 25,000 will be hired and trained for short 6 to 10 week assignments.

III. Testing the Increased Use of Administrative Data

10. In 2014, Statistics Canada launched a research program to study the potential to increase the use of administrative data on the Census of Population. One of the objectives of this research is to see if data from administrative sources could be used to replace the traditional collection of the census head count and of basic demographic characteristics. Canada uses a more traditional approach for the Census of population as it does not have universal population or dwelling registers, and no universal unique personal identifiers (Royce, 2013).

11. As one of the main initiatives to support the research on administrative sources, Statistics Canada has been building prototypes of the Canadian Statistical Demographic Database - CSDD (Cyr and al., 2013). This database is being built starting with federal income tax files, adding births and subtracting deaths, and accounting for some sub-populations via alternatives files such as the Indian Register and the Immigration

longitudinal Datafile. The first 2 prototypes of the CSDD were built to represent the population as the reference year of 2011 and produced populations that are within 1% of the national population estimates for the same period. But the prototypes demonstrated larger gaps with estimates at lower levels of geography, or for some segments of the population as defined by age and sex.

12. Statistics Canada has not defined yet what the criteria for success nor the timelines should be for completing this work. In the interim, possible set outcomes of the CSDD will none the less be evaluated over the next 2 years. These may not require the same level of global coverage to serve the evaluation purpose. For example, if timely basic demographic information was available for a large number of dwellings, this information could potentially be used to replace expensive field follow-ups during census operations. The information could possibly be used to validate dwelling occupancy status on Census Day, replace part of the field non-response follow-up, target difficult to enumerate areas, support work on coverage studies, etc.

13. Creating a prototype of a population spine (CSDD) has a number of challenges. One is timeliness. Can a file containing a large portion of the population be created in a timely fashion to coincide with census operations? The challenge here is the time necessary to access and match files of the proper vintage, i.e. the information must be as up-to-date as possible. Another challenge in the Canadian context is the ability to put people found on administrative files in the proper dwelling, i.e. at their primary residence. Most administrative files available to Statistics Canada do not require this type of address information. Address information may relate to secondary residences, or in the case of tax data, it could even the address of the accounted who completed the tax form.

14. Two sets of tests are being contemplated to test the application of the CSDD to support census operations. Potential impacts on data quality of replacing direct enumeration with administrative data could be verified by simulating the use of the CSDD in an operational setting with 2016 Census information. Control tests will be designed to see what would have been the impact of using the CSDD instead of continuing non-response follow-up on some dwellings in 2016. For example, we could institute a rule where no more follow-ups would be conducted on specific dwellings after ‘x’ attempts have been made. A pre-condition for this rule would be that good information on the occupants of the selected dwellings for the test exists in the CSDD. The vintage of the CSDD would have to be as close to Census Day 2016 as possible (i.e. May 10). The test would verify that direct collection at these dwellings can be replaced by information on the CSDD without important impacts on the data results at the small area level. Since the CSDD would not contain all variables collected on the census questionnaire (information on language would be missing for example), the missing information would be imputed. The results on the new file would then be compared to the official Census results from 2016 to verify accuracy.

15. One of the main benefits of substituting field follow-up with administrative data would be large cost savings. Such cost savings would be realised by reducing the number of contacts with non-responding dwellings, one of the most expensive operations during the Census. Part of the cost savings would be related to fixed costs in relation to the more rapid downsizing of the field infrastructure as collection is stopped earlier in some areas.

IV. Testing the collection of a personal identifier

16. All record linkages performed as part of the administrative data research at Statistics Canada are based on combinations of respondent names, age or date of birth, gender and address. Although there are no universal unique personal identifiers in Canada that could

be used to facilitate these linkages, the Social Insurance Number (SIN) could be a valid substitute. Most Canadians have a SIN and this number is also available on some of the administrative files available to Statistics Canada, such as income tax records. The challenge is that Statistics Canada does not currently have readily access to SIN information.

17. The collection of Social Insurance Numbers directly from respondents was tested as part of BT2. This portion of the test aimed at verifying the willingness of Canadians to provide their SIN on the census questionnaire, their ability to provide accurate numbers in a proxy mode (one person usually completes the census questionnaire for the entire household), and potential impacts on rates of response and return for both the Census and National Household Survey because of privacy or other concerns. These affects were measured for both the internet and paper modes of response and measured against control panels where the SIN was not collected. The Census portion of the test was also conducted under the mandatory provisions of the Statistics Act to replicate conditions that would actually exist on the Census.

18. The test showed that the majority of people would provide their SIN and that quality would be mostly high. Just over 80% of respondents to the test provided a SIN, and only 1.6% provided an invalid number. The analysis of the results did not focus on respondents who do not have a SIN so the actual proportion of respondents who do not have a SIN is not known. Some operational impacts were also observed during the test. Respondent burden increased by at least 20% for questionnaires with the SIN question as measured by the extra time required on internet to complete the Census questionnaire. Decreases in return rates were also observed – about 2% for the National Household Survey and 2% to 3% for Census at waves 2 and 3 (when reminders are sent to non-respondents). The slower rate of return does not necessarily lead to a lower response rate at the end of collection, but would mean a larger workload at the start of non-response follow-up, likely increasing costs for field collection.

19. Further analysis demonstrated that having the SIN did not significantly improve linkage rates to tax records (less than 1% increase), validating that the current approach used at Statistics Canada is fairly efficient. Having the SIN though would obviously speed up the linkage process. Based on the inconclusive results of the test, the SIN will not be collected as part of the 2016 Census.

V. Live Tests Planned for 2016 Census Operations

20. Statistics Canada will be conducting a few tests during the collection operations of the 2016 Census. These tests will be conducted to generate knowledge to support planning for the next Census in 2021. They include the collection of an e-mail address as contact information, the collection of both mailing and listing addresses if different, the use of digital delivery for the invitation to complete the census, and the extension of mail-out in certain areas.

21. Census respondents will be asked for their e-mail address on the 2016 Census questionnaire. This information will be added on the Household Survey Frame which is used as the source of addresses for the Census. The collection of the e-mail address is part of research on alternative modes of contacting Canadian households, either for the Census itself, or for any survey using the Household Survey Frame as a sampling base. The test will measure compliance (what proportion provide an email address), the capacity to accurately collect this information, and the possible implications on data capture operations for those responding on paper questionnaires. More particularly on the last aspect, we will

verify to what extent the optical recognition process is able to directly read and recognize e-mail addresses.

22. Parallel to the collection of the e-mail address, Statistics Canada will be testing an electronic approach for contacting households during collection operations in 2016. The test will involve contacting 50,000 to 100,000 households using the digital mailbox service offered by Canada Post (epost). This approach will be used on a test panel of epost subscribers and replace the planned standard mail contact. We will also be selecting a control panel from epost subscribers to validate the results from the test panel. The epost service will require households to first agree to have Statistics Canada as an authorized organisation to send material to their digital box. For those who do agree, they will receive their invitation to complete the census via their digital mailbox. The content of the digital message will be very similar of the paper invitation letter sent to other households via traditional mail. The test will measure rates of return and response of the digital group compared to the control group. The results will be used to plan household contact strategies for 2021.

23. Statistics Canada is using a number of strategies to contact households for the Census. Current assumptions for 2016 are that approximately 82% of dwellings will initially be mailed an invitation letter, inviting the occupants to complete the Census online. For about 17% of dwellings, Statistics Canada will deliver the questionnaires directly to dwellings in a 'list and leave' operation, primarily in rural areas, as the mailing system does not make possible to deliver unaddressed mail to specific dwelling. Residents of these areas receive their mail in common mail boxes, and since Statistics Canada can only address letters to 'occupants'.

24. The delineation of mail-out and list/leave areas is partly done considering operational criteria. For example, some dwellings which could be mailed to remain in list/leave areas to avoid creating mixed methodology areas for field staff, simplifying procedures. For the proposed test, we will test a 'mail-out with top-up' approach. For some areas where Statistics Canada has a large proportion of valid addresses, we will be mailing an invitation letter to these addresses, and a list of remaining known dwellings for which we do not have a valid mailing address will be provided to field personnel for the letters to be dropped off. The test will verify that this approach can be implemented, that costs can be reduced in comparison to a pure list/leave approach, and that dwelling coverage can be maintained. If the approach is successful, Statistics Canada could possibly expend mail out to at least 90% of dwellings in 2021.

25. Finally, Statistics Canada will also be collecting both the list address and the mailing address for each dwelling. The Address Register contains list address information, which sometimes does not match the mailing address, especially in rural areas. Obtaining concordance between these two addresses may allow the expansion of the mail-out universe in the next census in 2021. It will also support the research on administrative sources by allowing to place people in the right location, a key component of building a population spine

VI. References

Boulet, C., Mathieu, P., Statistics Canada, 2015, Testing the addition of the Social Insurance Number to the Canadian Census of Population; Presented at the Joint Statistical Meetings of the American Statistical Association, Seattle, August 2015

Royce, D. 2011. Preliminary Report on Methodology Options for the 2016 Census. Prepared for the 2016 Census Strategy, Census Management Office,

Cyr, A., Pignal, J. Trépanier, J., Statistics Canada, 2013 Canadian Statistical Demographic Database: A Research Project, Presented at the Statistics Canada Methodology Symposium