

**CONFERENCE OF EUROPEAN STATISTICIANS**

Joint ECE-EUROSTAT Work Session on Population and Housing Censuses  
(Ohrid, The former Yugoslav Republic of Macedonia, 21-23 May 2003)

**Session I – Invited paper**

**2001 CANADIAN CENSUS EXPERIENCE: SHIFTING GEARS**

Submitted by the Statistics Canada <sup>1</sup>

**I. INTRODUCTION**

1. Conducted every five years, the Canadian Census of Population is a major undertaking whose planning and implementation spans a period of over seven years. Statistics Canada generally works on two and even three censuses at any given point in time. Before the final results of one census are out, planning and systems development are already well under way for the next one. In this environment, introducing major changes to content or to operations from one census to the next is a major challenge. At the same time, the census in Canada can no longer remain as it has for the last thirty years, the last time a major operational change was introduced. Both external and internal pressures have made it necessary for Statistics Canada to consider collecting, processing and disseminating census information a different way.

**II. BACKGROUND ON THE CANADIAN CENSUS**

2. On May 15, 2001, Statistics Canada conducted its latest census, as over 46,000 field staff delivered census forms to about 13 million households. About 98% of the households were enumerated using the self-enumeration/mail-back method. An adult was asked to complete the questionnaire for themselves and for members of their household and return it by mail.

3. About 2% of households were enumerated using the canvasser method, where an enumerator visited the household and completed the questionnaire by interview. This method was used in remote and northern areas of the country and on most Indian reserves. It was also used in large urban downtown areas where residents are transient.

4. Most households (80%) received a short census questionnaire which contained seven questions on basic topics such as relationship to person 1, age, sex, marital status, and mother tongue. One in five households (20%) received a long questionnaire which contained the seven questions from the short form plus 52 additional questions on topics such as education, ethnicity, mobility, income and employment.

---

<sup>1</sup> Paper prepared Rosemary Bender.

### III. DETERMINING 2001 CENSUS CONTENT

5. Various social policy issues influence the content of the census in Canada. For example, recent censuses consider the demands of an aging population for necessities such as medical services and housing (age, sex and marital status questions); Canada's programs and policies related to multiculturalism, bilingualism - Canada's two official languages are English and French, and employment equity (immigration, language, citizenship and ethnicity questions); changes in technology require upgrading of skills and continued learning (education questions); the prevention of economic inequality (income, education and labour force questions); and the demand for affordable housing programs and establishing "core housing need" (dwelling and housing questions).

6. Although the 2001 Census questions covered essentially the same subjects as in 1996, changes were made to keep pace with the ever-evolving face of Canadian society. Statistics Canada undertook a rigorous consultation, testing, review and approval process to ensure that the 2001 Census addressed the highest priority information requirements of data users and the general public. The data needs were evaluated in light of a number of considerations, always bearing in mind the need to respect the respondent's right to privacy. These included legislative requirements, availability of other sources to meet data needs, respondent burden, and collection and processing costs. Proposed census content was rigorously tested using a variety of methods including qualitative testing such as focus groups and in-depth interviews with respondents and more formal national statistical tests.

7. This process presented two major challenges for Statistics Canada. The first was one of timing given its short five-year census cycle. The 2001 Census consultative process took place between autumn 1997 and spring 1998, before many of the results from the 1996 Census were made available. This early consultation was necessary in order to develop questionnaires for the national statistical test conducted in autumn 1998. The Agency addressed this challenge by taking into account feedback from multiple sources available up to that point in time. This included ongoing feedback from users of census data; ongoing consultation with major government census stakeholders; and the previous consultative process for the 1996 Census. As content was being developed for 2001, feedback from the current consultative process was considered, though at one point it became difficult to entertain suggestions for major content changes.

8. The second challenge was to manage stakeholders' expectations. Major changes to the census content generally occur over a number of census cycles. They are also often supported by changes to government legislation or programs. Modifications to question wording or response categories on the other hand are more feasible within a census cycle despite the short time frame. Underlying all discussions is the impact of proposed new content on historical comparability of census data as well as respondent burden, cost and operational considerations. It was thus important to set up a framework whereby stakeholders were able to present their points of discussion within the census context. To this end, Statistics Canada prepared a 2001 Census consultation guide which explained these considerations and for reference included the reasons questions were asked in previous censuses, related major social policy issues and alternative sources of data for the variables in question. Participants in the consultation process were asked to prioritize their suggestions and provide the underlying rationale. In this way Statistics Canada has been able to promote all elements of its socio-economic statistics program to users, not just the census.

9. Statistics Canada used the results of the 1998 national census test together with the results of focus group discussions, recommendations from major data users and national and community organizations and letters from private citizens to develop options for consideration and decision by federal Cabinet. As a result, the 2001 Census included new questions on language at work and birthplace of parents. As well, in view of the legal recognition of same-sex couples and the related need by governments and the private sector, the definition of common-law couples was changed to

include same-sex couples. Of the two questions traditionally asked every ten years, religion was included once again for 2001 while fertility was dropped.

#### **IV. 2001 CENSUS COLLECTION AND PROCESSING**

10. The census collection process in Canada is a highly decentralized, manual collection operation involving a large and geographically dispersed workforce. During the week of May 1, 2001, over 46,000 field staff including 34,000 enumerators delivered census questionnaires to 13 million households. Hired for six weeks, the enumerators were tasked with ensuring that all dwellings in their designated area received a questionnaire. Respondents were asked to complete the questionnaire and return it by mail. Questionnaires returned by mail were, in turn, forwarded back to the enumerator for editing. For each incomplete questionnaire, the enumerator followed up by telephone and/or in person to obtain the necessary information.

11. The recruitment and retention of tens of thousands of persons for a short six-week period is becoming increasingly difficult. The healthy economic situation in Canada in 2001 made it even more of a challenge to attract and retain census enumerators. The dedication of census staff and sound management practices in the end ensured successful completion of the collection exercise.

12. About one month following the collection of 2001 Census data, Statistics Canada began translating about 13 million questionnaires, or 5 billion keystrokes, into meaningful data. The first processing phase has essentially remained unchanged since 1981, as regional processing staff manually prepared the completed census questionnaires for data entry. For the fifth consecutive census, the capture of census data via key entry was contracted to Canada's national revenue agency who have been able to provide a readily available, experienced keying capacity at a reasonable cost. Scanning technology was introduced for the first time in 2001 to facilitate the verification and correction of the data as discussed further.

13. Three evaluation studies were undertaken following the 2001 Census to measure how well the population was counted. The Automated Match Study detected households and persons that were counted more than once in the same region by automatically matching samples of pairs of households with similar demographic characteristics. The Collective Dwelling Study also identified census overcoverage by obtaining an alternative address for a sample of persons enumerated in non-institutional collective dwellings. The largest evaluation study is the Reverse Record Check, where a sample of persons was selected from a list of persons who should have been enumerated in the 2001 Census. The list was created from the 1996 Census and administrative data sources detailing population change since then including births, immigrants and non-permanent residents. A sample of persons from the list were contacted to obtain their address on Census Day and any other address where they may have been enumerated. Preliminary results from these studies show a small increase in net census undercoverage for the 2001 Census compared to 1996. Net undercoverage continues to be highest for young adults, particularly young men. Undercoverage was also higher for the northern regions and in provinces with high immigrant populations.

#### **V. DRIVERS FOR CHANGE**

14. The census collection methodology in Canada has remained essentially unchanged since it was introduced in the 1971 Census. Since the early 1990's it has become increasingly evident to Statistics Canada that a change from these traditional approaches was required.

15. One of the original drivers for considering a change was the issue of confidentiality (in particular concerns with local enumerators) and security of personal information. These are increasingly of concern to Canadians, whether it be with regard to public or private institutions. While Statistics Canada offers the assurance of both confidentiality and security guaranteed under federal legislation, complaints from respondents are a source of concern. Under the current

collection methodology, completed questionnaires are returned to the local enumerator for completeness checks and follow-up when necessary. Some respondents object to local census representatives who may be neighbours or acquaintances seeing their personal information. The physical security of completed questionnaires is also of concern to Statistics Canada. They are typically kept in the enumerator's home or vehicle for some period of time until the assignment is completed. This leaves them vulnerable to loss or theft. Although to date the number of formal complaints in Canada arising from respondents is small and for the most part addressed by the field staff, there is the potential for such problems to become much more serious. This would certainly have a negative impact on future census collection activities.

16. A second driver is the federal revenue agency now known as Canada Customs and Revenue Agency. Since 1981 Statistics Canada has relied on its staff and facilities to key enter data from the census questionnaires. CCRA is moving away from its own key data entry operations as it aggressively promotes electronic filing of tax returns. They now receive well over 40% of tax returns electronically each year, either directly from the tax filer or through companies that offer on line tax preparation services. This percentage increases each year. As a result Statistics Canada has been examining alternatives to manual key data entry.

17. A third driver for change is the federal government initiative where by 2004 all transactions that Canadians have with the federal government should be available on line. Statistics Canada has already made significant inroads in the dissemination of census information to Canadians via the Internet. The next step is to provide Canadians with an option to fill out the census questionnaire on line.

18. In response to these drivers for change, more centralized and automated collection and processing methodologies are being developed. They consist of mailing out census questionnaires to respondents who would then return the completed forms by mail to a highly secure data capture centre. Data entry and completeness checks are being automated and any follow-up would be conducted by telephone from Regional Census Centres. This will eliminate many of the confidentiality and security risks found in the current methodology. It will also permit earlier processing and result in more timely release of census data.

19. A full scale centralized mail out/mail back collection and follow-up methodology is planned for the 2006 Census. Pilot tests focusing on certain key aspects of this new approach were conducted during the 1996 Census and the 1998 national census test. The cornerstone to this new methodology will be a high quality address register. The current planning assumption for the 2006 Census is that the address register will be used for at least two thirds of all Canadian households. The traditional method with enumerators will be used in those areas not well covered by the address register.

20. Statistics Canada is also looking to provide respondents with an option for filling out the census questionnaire on line. Many Canadians already use the Internet to file their tax returns, do their banking or purchase goods and their numbers are increasing all the time. The 2001 Census included a small pilot to test the logistics of an Internet option. Statistics Canada obtained valuable feedback on the process and plans to introduce this option to all Canadians for the 2006 Census. By then it is thought that 60% to 70% of Canadian households will have access to the Internet. With a more aggressive communications campaign it is expected that many of these households will fill out the questionnaire on line.

21. The change in collection methodology for 2006 cannot be undertaken without parallel and significant changes in the capture and processing of the census data. The challenges of replacing what is essentially a paper processing operation with a virtually automated system are enormous, not the least of which are the significant costs, time and effort associated with building major systems applications. There is also the added complexity of integrating this with the new centralized mail out/mail back collection methodology.

22. Recent years have seen the development of data capture technologies well suited for census processing applications. Statistics Canada has opted to introduce imaging technology in phases. For the 2001 Census, images were used primarily for the verification and correction of data rather than for a full integration of data capture and processing operations. It also permitted the census to obtain experience with the new technology with relatively low risk to the output. For the 2006 Census, Statistics Canada has decided to contract outside the Agency the development of much of the automated processing operations. This includes the development of scanning and autocoding applications, computer telephone applications for assisting with direct response by phone and with the field edit follow-up, as well as for the integration of information collected by mail, on line and by telephone. There is not sufficient time between quinquennial censuses for Statistics Canada staff to learn, develop and apply the necessary expertise. Rather, the Agency is drawing upon international expertise in the industry as other countries are also adopting these technologies for their censuses.

23. Thus for the 2006 Census, the collection and processing systems will be fully integrated as key entry activities will be replaced totally by scanning and automated data capture operations. In the long run this should result in efficiency and timeliness gains, as well as address respondent concerns over confidentiality.

## **VI. COMMUNICATION AND DISSEMINATION OF 2001 CENSUS RESULTS**

24. The 2001 Census data were released in eight phases. Population and dwelling counts were first released in March 2002. Demographic and socio-economic characteristics of the population, of families and of households were then released over the following fourteen months with the last release in May 2003.

25. Major changes to the approach for releasing the 2001 Census data had a significant impact on the ensuing coverage of the results. During the planning stages, Statistics Canada laid out two specific objectives to enhance the communication of census results. The first was to provide more relevant analyses of the census data. To this end, storylines for the analysis focused on issues and were prepared months in advance. In anticipation of the 2001 data, analysts initially used data from previous censuses as well from other surveys. The resulting 2001 Census Analysis Series, written for the general Canadian public, provides a comprehensive socio-demographic portrait of Canada at national, provincial, territorial and regional levels. The Internet versions of these reports link to data tables, charts, maps and flash graphics.

26. The second objective for the Agency was to improve how it conveyed information to Canadians, particularly about their own community. Highlight Tables were created with demographic and socio-economic indicators providing for the first time measurements of their distribution and percentage change over time, which can be ranked and sorted across a wide range of geographies. Statistics Canada also worked extensively with the media weeks in advance to advise them of the census releases, without of course divulging 2001 Census data or trends.

27. As a result of these initiatives, Statistics Canada has exceeded its expectations. Interest across the country in the 2001 Census results is unprecedented, with extensive in-depth media coverage, record numbers of visits to the Statistics Canada Internet site and a large demand to speak at interdepartmental meetings and professional conferences.

28. While many of the operational changes introduced so far with the 2001 Census are transparent to the respondent, the same cannot be said for the users of the 2001 Census output. With this census, the Internet has now become the main vehicle for the communication and dissemination of census data. This preference was evident from consultation sessions conducted by Statistics Canada with a wide range of data users, including federal, provincial and territorial government departments, municipalities, academics, librarians, the private sector and secondary distributors. Though their data requirements vary considerably in terms of content and level of

geographic and demographic detail, a large number of them pointed to the Internet as a major mode of access.

29. The Census module on the Statistics Canada web site provides direct access to a wide range of products under the general headings of analysis, maps, data and reference products. The 2001 Census Analysis Series is housed under the first heading. Under maps the Internet user will find thematic maps complementing the analytic series as well as thousands of more general reference maps. A geographic mapping tool includes features that permit one to search and zoom in on geographic areas as fine as city blocks. The data products include three sources of data available at no cost to data users. A product called Community Profiles provides basic demographic and socio-economic information for all communities in Canada. The Highlight Tables mentioned earlier present small area data from an analytical perspective. A third source is the Topic-based Tabulations series that presents hundreds of cross sectional tabulations. More complex tables in this last series are available to certain stakeholders using passwords while still other tables are available to all data users for a fee. This structure has enabled Statistics Canada to efficiently meet the various data requirements of its users via the same platform. Finally, extensive reference material is available including data dictionaries, users guides, geographic name changes and data quality notes.

## **VII. CONCLUDING REMARKS**

30. Statistics Canada is addressing both internal and external pressures to change its collection, processing and dissemination strategies. Using a number of pilot tests and phasing changes over the 2001 and 2006 Census cycles, the Agency is moving away from a decentralized, manually intensive collection and data entry operation to a more centralized and automated approach. This will in particular address key concerns regarding confidentiality and security of personal census data. At the same time, a more proactive census communications and dissemination strategy has been adopted which has led to substantial increases in media coverage and interest by a large number of census stakeholders.