



**Economic and Social
Council**

Distr.
GENERAL

ECE/CES/GE.41/2007/9
22 March 2007

Original: ENGLISH

ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Group of Experts on Population and Housing Censuses

Tenth session

Astana, 4-6 June 2007

Item 3 (b) of the provisional agenda

**CENSUS TECHNOLOGY: RECENT DEVELOPMENTS AND IMPLICATIONS ON
CENSUS METHODOLOGY**

Census on the Net

Submitted by Canada*

This meeting is organised jointly with Eurostat.

Summary

The Bureau of the Conference of European Statisticians (CES), at its meeting held in Washington, D.C. (United States) on 19-20 October 2006, approved the renewed terms of reference for the Steering Group on Population and Housing Censuses and the plan for future CES activities on population and housing censuses. The CES Bureau also agreed that the Steering Group would coordinate the work on the diverse types of meetings. The present paper was prepared on request by the Steering Group on Population and Housing Censuses, for presentation and discussion at the Joint UNECE/Eurostat Meeting on Population and Housing Censuses in Astana (Kazakhstan), 4-6 June 2007. The paper provides substantive basis for the discussion in the session of the meeting dedicated to "Census technology: recent developments and implications on census methodology".

* The paper has been prepared by Statistics Canada.

I. INTRODUCTION

1. Statistics Canada, like several other statistical agencies around the globe, provided a secure online Internet based application so that everyone who lives in Canada could complete their 2006 Census questionnaire online. This short document will highlight some of what Statistics Canada have learned following the decision to offer an online option for 2006 and testing it in a sizeable dress rehearsal in May 2004. The introduction of this added response channel is one of the major changes to census methodology in Canada in more than 30 years. Specifically, this paper examines basic pre-conditions that should exist prior to the consideration of offering a census online option, strategic and business case considerations, main risks, challenges and benefits, and long-term opportunities.

II. BACKGROUND

2. The 2006 Census provided the capacity for respondents in all private households and agricultural operations in Canada to complete either the long (53 questions) or short (8 questions) questionnaire by Internet. It was anticipated that approximate 16 per cent of respondents would use the online application, which was available 24 hours a day, 7 days a week. A comprehensive public communications strategy was planned to encourage respondents to complete their census form online. Census questionnaires were mailed to about 73 per cent of the dwellings in Canada, with the remainder hand-delivered by an enumerator. A unique Internet access code was pre-printed on the front of each paper questionnaire. Respondents completed and transmitted their questionnaire directly to Statistics Canada's census Data Processing Centre, where the response was registered and its data integrated into the regular flow of census returns. Responses received on paper and through electronic channels were logged against a master control list of dwellings in order to track the status of all dwellings. This was essential in order to allow notification to field staff to either conduct or suspend non-response follow-up.

3. The Internet application was designed so that after a respondent uses it to submit their questionnaire, no software trace (footprint) was left on his/her computer. Respondents who completed the long census questionnaires were able to save their partially completed form and establish a password. By using their password in conjunction with their Internet access code, respondents could complete their form over multiple sessions. If a respondent chose to save their form and did not return within a pre-specified period of time, the partially completed form was submitted on their behalf automatically, and was passed on to subsequent processing steps. In addition, Internet respondents who saved their form could access it from different locations (roaming). For example, they could start the form at home, save their partially completed form and complete it from work. The Internet application was available in both of Canada's official languages (English and French), and respondents could toggle back and forth between languages as they completed the form. Once the on-line questionnaire was submitted, the subsequent processes were independent of the response mode, be it through paper or computer assisted telephone interviewing, thus making it a truly integrated approach.

III. PRECONDITIONS

4. In the Canadian experience, there are a few factors that should be considered essential, before a census Internet option is offered to a country's population. In addition, there are a few key elements, which if present, would greatly enhance the likelihood of a relatively successful census online offering.

A. A methodology that is conducive to adding an Internet option

5. Adequate literacy levels and the ability of a large portion of the population to self-enumerate are absolutely essential for an Internet census option. Although the manner in which the questionnaire is delivered (manually or through the postal system) does not factor heavily, the ability to provide each household (or individual) with a unique Internet access code linked to a geographical location is key to having an effective tracking and control system. An exclusive or majority canvasser (face to face interview) methodology would certainly pose some significant challenges for the viability of an online option. For the 2006 Census, Statistics Canada created a list (the Master Control System or MCS) of all dwellings with a unique identifier linked to a physical address (where a questionnaire was mailed-out), and linked to a geographical code (where a questionnaire was manually dropped off). Each questionnaire was pre-printed with this unique identifier as well as a unique Internet access code. When questionnaires were mailed back, received via the Internet or via the Computer Assisted Telephone Interview application (CATI), they were registered using this unique identifier and the MCS was updated. Non-response follow-up lists were created from the MCS and transmitted to Field staff on a regular basis, allowing contact with respondents only for those dwellings that had not responded.

B. Connectivity and public expectations

6. Clearly any country considering the viability of developing an application for online filing of census forms must consider the level of connectivity amongst the population, their access to high-speed Internet, as well as the extent and ease with which the population conducts "business" using the Internet (for example, online banking, filing tax forms). The latter may be indicative of the level of trust of online applications among the population and a potential indicator of the willingness of respondents to provide their confidential census information online.

7. In Canada, a majority of households are connected via the Internet. Based on the 2003 Household Internet Use Survey, an estimated 7.9 million (64 per cent) of the 12.3 million households in Canada had at least one member who used the Internet regularly in 2003, either from home, work, school, public library, or other location. An estimated 6.7 million households (54 per cent) had a regular user from home, with 4.4 million (65 per cent of the home users) with high-speed access to the Internet. Furthermore, about 57 per cent of households using the Internet at home had someone who accessed online banking services, well above the 44 per cent in 2001. This growth indicates that Canadians are becoming more confident in the security aspects of the Internet.

8. High in the consideration to offer an online option was the fact that the enormous penetration of the Internet in Canadian households, along with the sophistication of its use in

differing types of transactions, would result in a strong expectation on the part of the public for a secure and efficient online option for the 2006 Census.

C. Mature security infrastructure

9. While Canadians in general are taking to the web, trust in the level of security offered by the vendor or institution plays a big factor in their decision to either conduct or not conduct an online transaction. For a statistical agency, the provision of a secure infrastructure thus becomes absolutely essential in ensuring that the trust factor is high enough for the respondent to be willing to provide confidential census data online. For Statistics Canada, given this first-time national offering, it was important to provide an infrastructure that was in reality as, if not, more secure than most other online transactions to which Canadians have become accustomed. The fundamental difference in a statistical survey or census is that the respondent is being solicited for a response, rather than initiating or benefiting directly from the transaction—such as in the case of online banking or requesting unemployment insurance.

10. Serious consideration should be given to the level and type of security as this affects the design and overall cost of the Internet option. There are a variety of options and undoubtedly more will emerge in the years to come. In Canada there was the choice of having a downloadable application that would have allowed the respondent to complete a form and would then encrypt the respondent's data before transmission. This was the small-scale model tested in 2001, and due to the lengthy download times and other technical factors, it was decided to pursue a different approach in 2006. Another choice was to use SSL128 that is being used by a number of banks. While this is fairly secure, STC felt that for the census it was required a higher level of security.

11. As part of its Government on line (GOL) strategy to have all services online by 2005, the Government of Canada pursued the creation of a "Secure Channel" through which citizens could transfer confidential information using the Internet. This initiative is offered to Federal departments as they implement their individual online strategies. Secure Channel's original plan was to only offer unidirectional encryption (i.e. encrypt the data flowing from the user's computer to the department's servers). Statistics Canada required bidirectional encryption since confidential information would also be flowing from its servers to the user's computer. This functionality was required in order to conduct online edits as well as to offer respondents the choice to save and resume a session in the future. Statistics Canada made the investment required to customize these requirements for the Census within the Secure Channel environment. This, with the use of anonymous, limited-use certificates, and the isolation of Statistics Canada's internal network brought an exceptionally high level of security for the information. Specifically, Statistics Canada wanted to ensure that there was no possibility, or a possible perception, that Canadians' census information could go to anyone other than Statistics Canada, and that the digital certificates were not "generic" but used exclusively for the census.

12. The scale and scope of the Census requires a robust security infrastructure, and while it was possible for Statistics Canada to develop it for the Census specifically, with the flexibility shown by the government wide initiative to accommodate the unique needs of Statistics Canada, it was deemed less risky and more cost-effective to leverage from the investment on this initiative.

D. Online surveys experience

13. Pre-2001 Census consultations indicated a growing demand and expectation for an online application for the 2001 Census. Statistics Canada met this demand by offering a relatively inefficient but secure and functional application to a very small sub-set of the Canadian population. The Internet option was made available in two specific test sites to approximately 190,000 households and 8,300 farms. The only public communication message was through an insert in the questionnaire package that also contained the unique access code for the respondent. Security requirements meant that the online application had to be downloaded to the respondent's computer. This turned out to be an onerous task as the application size was 6 to 7 MB, representing a download time of approximately 30 minutes with a 56K modem. A total of 4,300 Internet responses were received. Lessons learned involved mainly design issues, as well as the recommendations for 2006 to offer a true online application that did not have to be downloaded and in fact, one that would leave no footprint on the user's computer.

14. Not offering an Internet option in 2006, given the expressed demand and expectation in 2001, would have been out of step with the government's GOL initiative. The lack of an Internet option could potentially have resulted in some non-response by specific groups expecting and demanding an online response option, posing unacceptable risk and costs for 2011. Statistics Canada leveraged private sector expertise in the development of the 2006 Census application/infrastructure. Despite leveraging from the private sector, it is beneficial if in-house experience and expertise can be implicated in defining, developing and testing user-requirements in a survey taking context.

15. Many of Statistics Canada's business surveys have offered respondents the option of providing data electronically. While these initiatives are largely geared to reducing respondent burden for companies who supply data on an on-going basis, and there are significant differences between the census and business surveys in terms of content and the respondent community, Statistics Canada have successfully leveraged from these initiatives to plan and implement the census infrastructure and to address a number of risks.

16. Thus prior to embarking upon an online census offering, factors such as a compatible collection methodology, an effective control-tracking system, high connectivity and public acceptability, and the presence of a mature security infrastructure are deemed essential. Previous online survey and/or Census experience is certainly considered to be a desirable asset in potentially reducing some of the risks associated with such an undertaking. Above all factors however, is the willingness of the statistical and any funding agencies to consider an Internet response channel as a potential long term investment with a clear understanding of the potential risks and opportunities. Sufficient funds must be clearly identified in advance and not predicated on uncertain future savings.

IV. STRATEGIC CONSIDERATIONS

17. Having worked through the essential and highly desirable pre-conditions for offering an online census option, a number of strategic considerations weigh heavily on the manner in which to proceed with the design, planning and implementation of such an undertaking.

A. Investments and potential savings

18. The main cost drivers for the Internet application are questionnaire development and testing, developing a secure infrastructure, the application development, testing (functional, integrated and volume) and support (technical and respondent) efforts, telecommunications infrastructure, hardware/software and public communications expenditures. In 2006 Census context, about half the expenditures were associated with the customization of the secure channel infrastructure and could possibly be considered as a one-time investment. The maximum number of concurrent users plays a critical role in determining the capacity and thus cost, and in this case, the infrastructure was sized to accommodate 15,000 concurrent user sessions. The total target response rate over the Internet was 20 per cent of households that respond prior to non-response follow-up (approximately 2 million questionnaires), which if not met would have a direct impact on the capacity of the paper-handling systems and operational costs. In 2006, prior to the start of the non-response follow-up, about 22 per cent of responses were received via the Internet. Given that the non-response follow-up stage is almost exclusively conducted with paper questionnaires, by the completion of data collection operations, some 2.26 million or 18.5 per cent of the total households in Canada used the Internet response mode.

19. While the targets were established for the entire Census collection period (some 9 weeks from start to finish), Statistics Canada had to specify and build technical capacity to sufficiently handle a finite number of sessions at any one time. With very little experience, and only a theoretical model on which to estimate infrastructure demand, Statistics Canada developed a system of close monitoring, oversight, and escalation and introduced the notion of “graceful deferral” to balance load, performance, cost and potential risk. The notion of graceful deferral, allowed a pre-determined number of respondents already on the system to continue uninterrupted at an acceptable level of performance, and prevented any additional users from accessing the system to the point where the overall infrastructure could no longer be sustained. Statistics Canada faced enormous demand on Census night, especially between 6 and 11 pm, and invoked graceful deferral repeatedly during this peak period. It is estimated that approximately 150,000 users were asked to re-try the application at another time.

20. The potential savings opportunities are in the areas of reduced paper-handling costs, mail-back postage costs, and reduced follow-up costs for missing or inconsistent information. For 2006, with a 16 per cent Internet return rate, these were anticipated to approximate the cost of developing and supporting the Internet application, and thus place in a near break-even position. The 2006 Census achieved and slightly surpassed this objective.

21. Based on the willingness of Canadians to use the on-line census application in 2006, Statistics Canada is in the process of developing planning assumptions for 2011, which maximize the benefits of on-line reporting to respondents and offer cost-savings for the government.

B. Developing the application in house or contract out?

22. In 2001, with the help of internal and external advisors, Statistics Canada developed and subsequently applied criteria to determine which components of the 2006 Census were best developed in-house versus by the private sector. One of the key considerations was the level of expertise and where this was best capitalized upon in order to minimize development, integration

and implementation risks. The conclusion was that Statistics Canada had the in-house expertise to design, user-test and specify the online questionnaire format, flow, skip-patterns, and rules for online edits. These factors influence the respondent experience, the overall look and feel of the application, comparability to the paper questionnaire (possible mode effects), as well as the resulting data quality and as such deemed core to the expertise of the agency. Conversely, it was concluded that the private sector was better capable of meeting the systems integration challenge associated with programming and linking the online application with downstream processing systems and ensuring that it was optimized to handle the anticipated volumes.

23. Thus conducting a competency profile, with a comprehensive view of the overall census system and process design, proved beneficial in capitalizing on the respective talents of the in-house staff as well as those of the private sector.

24. Having made the distinction at a high-level, there are grey areas where competencies overlap, and Statistics Canada worked with external contractors in developing the census Internet application in an integrated team environment. This type of working relationship yielded dividends even in areas where typically Statistics Canada has the upper hand. Statistical agencies are by far the best qualified when it comes to creating paper questionnaires. Decades of experience, huge samples, exhaustive data analysis and qualitative tests and studies have resulted in a level of expertise that is without equal in the private sector. Our experience in creating web based data collection vehicles is more limited and in some cases there is a risk that in attempting to adhere too closely to the existing paper form in an attempt to minimize a mode effect, it can actually be created a larger mode affect. Respondents who choose to respond via the Internet tend to be experienced Internet users, and have expectations about how the form should work based on their experience with other web-based forms. Working closely with the contractors as well as through the census test experience, useful lessons have surfaced on how to better design the web based census questionnaire to achieve desired results. For example, web based forms generally order yes/no questions consistently. On paper, it is tended to place the response that is most often selected first. On the web, users apparently learn the order of the yes/no responses from their first encounter with this type of question and then expect all others to be ordered in the same way. Switching the order of these responses on the web to keep them consistent with the paper form can result in the opposite of the desired effect.

25. Another example pertains to questions, for which there is only one logical answer (e.g. male/female) that use radio buttons. If the user selects "male" by accident, then selects "female", "male" is automatically deselected. An experienced Internet user expects this behaviour and a situation where the response ends up with both "male" and female" being selected is more likely if radio buttons are not used in the design of the questionnaire. In Canada, the potential implications were carefully considered before introducing anything in the Internet form that was different from the paper form, and for 2006, a balance was achieved with general adherence to the paper form, while incorporating many useful web standards.

26. Finally, despite the stringent security measures in place to protect confidential data, careful consideration must be given to how respondents may react to the private sector being involved with real or perceived concerns about data confidentiality and privacy. Any widespread negative reaction could have an adverse impact on overall response rates and create respondent relations issues for the statistical agency.

V. MAIN RISKS AND CHALLENGES

A. Integration

27. One of the major integration challenges was determining how the data received from the Internet would be integrated with those received from different streams such as paper and the telephone. Our decision to implement some of the web standards resulted in a small number of differences between the data received from the Internet and those received from paper. This required co-operation between those creating the database, and those designing the paper data capture system and the Internet questionnaire at all stages of the development cycle.

28. In trying to gain the maximum return on the Internet application development effort, a conscious design decision was made early in the process to try and use the same basic logic and code-set with only necessary modifications for two other similar applications. Specifically, in this context questionnaires were completed over the telephone in a Computer Assisted Telephone Interviewing (CATI) mode and a similar application was used for following up on failed-edit follow-up (FEFU) cases, once again using a CATI application. All three of these (Internet, CATI and FEFU) shared the same basic code-set with some customization. While there was some success with this design, using this approach resulted in a number of challenges. For example, while the Internet application had built-in edit-logic, the FEFU application had to also deal with data supplied on paper questionnaires where skips may not have been followed and inappropriate multiple responses may have been provided. In the FEFU application, it was necessary for the operator to address specific discrepancies with the respondent, by being able to access responses that didn't respect the logic built into the Internet application. These types of issues posed some challenges when scaling the FEFU application up for the 2006 census. The current plan is that the FEFU and Internet applications will not share the same basic code-set for 2011.

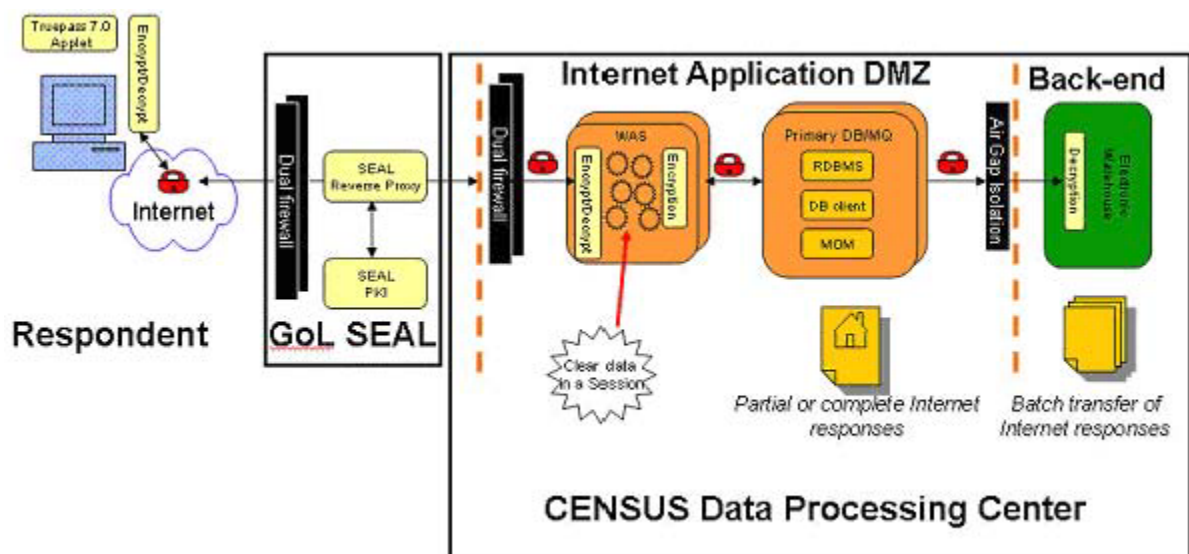
B. Infrastructure

29. One of the major challenges in implementing an online option for the census was the size of the infrastructure that was necessary to handle the peak number of users. The infrastructure generally consists of servers, storage devices and communication vehicles that can be extremely expensive to purchase, set-up and test. The strategy was to reach a balance in terms of what was a reasonable cost versus the risk of not being able to handle the peak number of concurrent users and having them give up on the Internet and complete paper questionnaires, or in the worst case scenario, give up entirely and not respond at all. The peak demand on the infrastructure is driven by 3 variables, the number of users logging in at any given moment during the peak time, the average number of pages each user will see, and the average length of time each page is viewed. For example, a less robust infrastructure is required to support a scenario where each of 50 users is presented one page that they will view for 5 minutes than is required for 15,000 users to view 20 pages for an average of 10 seconds per page.

30. Another part of the calculation to determine the infrastructure necessary is the total number of online forms expected. However, while users logging on and submitting forms represented the most significant burden on the infrastructure, other types of activities that consume system resources also require consideration. These include users who attempt to log in but who use invalid access codes, users who stop and save their forms to return another day and

those who log in, start their form but never save or submit it. Assumptions used to estimate the infrastructure required must include estimates for these activities in addition to the number of forms that will be submitted.

31. Another major consideration related to the infrastructure is the system's architecture, how it interfaces with the other components of the processing system, and how it secures respondent's data. STC's processing systems have no connections to outside networks. The Internet response channel (IRC) is obviously connected to the Internet. The STC architecture protects respondent data inside the IRC servers with a dual set of firewalls and by encrypting respondent data. It also moves the encrypted data from submitted forms to the main processing systems through the use of an "air gap" device that physically separates the two networks but allows authorized data to pass through a physical switch at pre-determined time intervals.



C. Communications

32. To have a reasonable prospect of success in achieving the desired uptake level, a consistent message denoting the ease of use of the Internet was developed and used in all materials, advertising and media interviews.

33. The following list provides some examples of how and where the one consistent Internet message was used:

- (a) All print and electronic material included a website address and the message "Complete your census form online today";
- (b) Graphics for the national census poster were focus group tested to ensure understanding of the image and message;
- (c) Paid radio advertising focused on the ease of completing the census form online. Print and out-of-home advertising included the envelope and mouse graphic;
- (d) All communications activities emphasized the Internet option for 2006;

- (e) Hundreds of public and private sector organizations agreed to promote the online option on their external websites during the six weeks prior to Census day, using a combination of information and hyperlinks to the online questionnaire;
- (f) Links to the online questionnaire were prominently featured on the Statistics Canada website;
- (g) The online questionnaire was emphasized repeatedly by staff in print, radio and television interviews.

34. Promoting the online option presented some unique challenges. For example:

- (a) Census Day saw an unexpectedly large volume of online activity. As a result, there were delays in accessing the form online;
- (b) Linux users were initially unable to use the online questionnaire. This technical issue was quickly resolved by STC.

VI. MAIN BENEFITS

A. Quality

35. A number of studies led to believe that the responses received from the Internet are more complete than those received from paper. In the census test, it was observed that the incidence of item non-response is many times lower for Internet submissions than for paper. Specifically, for short census of population forms the item non-response rate was 0.01 per cent for Internet responses and 2.54 per cent for paper responses. For long census of population forms the item non-response rate was 1.80 per cent for Internet responses and 6.97 per cent for paper responses.

36. Some of this can be explained by the fact that average Internet users tend to submit more complete responses than on paper. To determine if the Internet option improved the already high quality of responses submitted by the Internet user community, the data individual respondents submitted to the 2004 test were compared to that which they submitted on paper for the 2001 census. When controlling for the fact that the 2004 test was voluntary and the census was mandatory, it was found that the questionnaires these users submitted on the Internet in 2004 were more complete and had lower item non-response than those they submitted in 2001.

37. There are a number of factors that contribute to the improved quality observed with online responses. By design, a number of online edits have been built into the application that prompt respondents when they have left questions blank or entered conflicting information. Automated skips have been programmed that guide respondents past questions that are not applicable, and questions have been personalized with respondents' names which reduces the likelihood that one person's responses will be inadvertently recorded in another's response area. Also, no data capture errors are introduced during processing. The use of radio buttons in certain questions eliminates the possibility of conflicting responses and the use of drop down menus where appropriate, assists users in providing appropriate responses. Finally, there is a general perception on the part of users, confirmed by a survey conducted with a sample of Internet respondents after the census test, that the Internet form is easy to use and quick to complete. This is likely at least part of the reason behind the observation that more respondents completed the

whole form online (in comparison to paper) rather than getting discouraged and quitting before the end.

38. Completing a complex (household, matrix format) questionnaire online remains as yet somewhat unique and possibly considered by some as a novelty. It is not clear if the same perceptions with regard to ease of use or quality results will hold as more and perhaps less sophisticated users are enticed to use the online application in future censuses.

B. High level of user satisfaction

39. Ensuring a high-level of satisfaction with the online experience is essential in retaining the respondent for the duration of the session, resulting in a completed questionnaire. User satisfaction is influenced by a number of factors, such as the ease with which the questionnaire is accessed, which is influenced by the minimum system requirements of the security solution and the complexity of the login process. Other factors include the absence of lengthy downloads or foot-prints, working with the default configuration of most users' machines, being compatible with common browsers, intelligent skip patterns, soft edits that permit the user to continue even if a particular response was not valid or provided, and the respondent's perception of the relative speed and ease of use of the application. After the 2004 Census test, Statistics Canada conducted a follow-up survey with a small sub-set of those who responded over the Internet to judge their level of satisfaction. Some of the relevant results are as follows:

- (a) the majority of the respondents (89 per cent) indicated that they completed their Census test questionnaire at home;
- (b) 79 per cent of the respondents indicated that they had a "high speed" connection;
- (c) when asked why they chose to complete their Census test questionnaire online, 52 per cent said because it was easier, 30 per cent said because it was faster, 18 per cent said because of personal preference, 16 per cent said because they did not have to mail it;
- (d) 95 per cent of the respondents rated their overall experience with the Census Internet application as favourable;
- (e) 88 per cent of respondents felt that the time it took to complete the Census Test questionnaire was acceptable, of which more were short-form respondents. Respondents who indicated that they did not have 'high speed' were more likely to say that it took too long;
- (f) 98 per cent of respondents indicated they would complete their Census questionnaire online in 2006;
- (g) 57 per cent of respondents were not at all concerned that the privacy and the confidentiality of their Census test questionnaire data were more at risk on the Internet, while 34 per cent were concerned and 8 per cent were very concerned;
- (h) when asked how secure it is to transmit personal information over the Internet, 73 per cent stated that it was secure, 13 per cent stated it was very secure and 10 per cent stated it was not secure at all.

40. Three follow-up surveys were conducted after the 2006 Census. The survey results will be available by May 2007.

C. Long term opportunities

41. Predicated on a number of basic assumptions such as the continued acceptance and use of the Internet by most citizens, the Internet option holds prospects for long-term cost savings, quality gains, and reduced respondent burden.

42. In geographic areas with a very high Internet penetration rate, it may be feasible to offer the online option as the default with a simple mailing of the URL and an access code, possibly significantly reducing the printing and delivery costs of the paper questionnaire. In Canada this may present proportionally higher savings, as in most cases Statistics Canada has to print and deliver a questionnaire in both English and in French. The challenge is in being able to deliver a paper questionnaire in the language of choice in a timely fashion to those who cannot or do not wish to use the online application (through a phone call to the help-line).

43. As part of the 2006 Census, a study was initiated to test a method aimed at promoting the Internet response. A model was developed to identify *a priori* areas that include a significant number of dwellings likely to answer the Census online. Households in this study, called the Push Strategy, received only a letter instead of a paper questionnaire. These households were asked to complete their questionnaire online. The letter also included a 1-800 telephone number, which respondents could call for information about the study or to request a paper questionnaire. A preliminary sample of 40,000 households in mail-out areas was selected for this study. This sample was split randomly into two groups of 20,000 households each in order to create a control group. The control sample received a paper questionnaire while the Push Strategy sample received the letter.

44. The two most important constraints of this study were the inability of directly identifying households with an Internet connection and minimizing the burden on data collection operations. In that respect, the sample was not concentrated but was rather spread across target areas in Canada. Using data from the 2004 Census test and from the 2001 Census, a model was developed to identify areas that included a significant number of dwellings likely to answer the Census online. The model was based on the variables that were the most statistically different between the Internet and other response modes. For example, variables such as age, level of education, work place activity, student status, children at home, income and multiple family households were among the most significant. These variables determined an Internet 'profile' of people and households. This profile was then applied to the 2001 census data to establish a score for each individual and then each household. Analysis of the concentration of households with the higher scores was performed in order to identify areas with higher proportion of households likely to respond online. Finally, 2001 identifications were converted to 2006 areas using geographical conversion files.

45. The method was quite effective since the Internet response rate of the Push sample was 2.6 times more than the control group and 3.4 times more than the general population. Looking toward the 2011 Census, there are plans on increasing mail-out from 73 per cent to 90 per cent of dwellings and expanding the use of the Push Strategy.

46. The Census experience indicated a much higher rate of item response and very low edit-failure rates for the Internet questionnaires. A higher take-up rate would concurrently result in lower follow-up and imputation rates, resulting in lower costs as well as increased quality. The

online application also holds prospects for presenting a wider range of options to select from for certain responses, such as those in the labour or education categories (occupations, industries, educational attainment), thus reducing manual coding effort.

47. With an ever-increasing diversity in the Canadian population, the Internet makes it possible to present the census questionnaires in multiple languages at modest additional cost. Even if the respondent chooses to use the paper questionnaire, the Internet can be used to provide translation of the questions in multiple languages. Overall, these could hold the prospect of increased response and lower follow-up costs. Similarly, the online application provides some real alternatives to people with disabilities such as those who are blind to self-enumerate rather than rely on an enumerator.

VII. CONCLUSION

48. Statistics Canada's experience with the development of an online Census application for the 2006 Census has been positive. The execution of a sizable and live Census test in 2004 provided an opportunity to exercise an innovative design with respondents in an integrated collection and processing environment. A number of factors favoured the introduction of the online census application for 2006, such as the relatively high connectivity rates and comfort with transacting over the web by a majority of the Canadian population. The census collection methodology of mailing-out/dropping off questionnaires for respondents to self-enumerate was also conducive to the introduction of the online application. An integrated processing design was developed whereby all modes of collection culminate in the same downstream editing, follow-up, and imputation systems and processes. The contracting approach yielded a competent and experienced vendor, well positioned to deliver a robust, integrated and user-friendly application. The in-house online survey experience allowed to test and specify relevant and priority requirements. The timely development of the Government of Canada's secure channel initiative and their flexibility in adapting it to suit Statistics Canada's needs aided significantly in presenting a secure infrastructure. A robust governance and project management structure was instituted in order to ensure that design changes and costs were closely monitored and tightly controlled. All this meant that Statistics Canada was well positioned to meet the expectations of Canadians for a secure, efficient and user-friendly online application in 2006. Statistics Canada will build upon this successful experience towards 2011.

* * * * *