Catalogue no. 11-522-XIE

Statistics Canada International Symposium Series - Proceedings

Symposium 2003: Challenges in Survey Taking for the Next Decade

2003

Statistics Canada Statistique Canada

Canada

Proceedings for Statistics Canada Symposium 2003
Challenges in Survey Taking for the Next Decade

2006 CENSUS OF CANADA: STRATEGIC DIRECTION AND ISSUES

Anil Arora[1]

ABSTRACT

After a relatively stable period spanning 30 years, the manner in which a Census is taken in Canada is being fundamentally changed, responding to external pressures and taking advantage of opportunities in improved data processing and data collection technologies. A substantial change in collection and processing methodology will significantly reduce privacy concerns expressed by Canadians, present an alternative to a declining data capture capacity, provide an on-line response option, and provide opportunities for long-term cost savings and decreased processing timeframes. This paper outlines the changes in collection and processing methodology planned for the 2006 Census in Canada, and focuses on three primary areas: Content changes, Internet data collection and Outsourcing

KEYWORDS: Census; Internet; Outsourcing; Privacy.

1. BACKGROUND

The last major redesign of the Canadian Census collection process was undertaken in the 1971 Census. In each subsequent census, enhancements have been made to various modules of the Census process. The following is a description of the collection methodology and data capture approach employed in 2001 for the Census of Population and the Census of Agriculture, which are conducted simultaneously.

Each census representative was assigned an area for which s/he was responsible for all enumeration activities. Beginning two weeks before Census Day, the Census Representative visited each dwelling in his/her assigned area, delivered a questionnaire package, and listed the dwelling address in a control document called a Visitation Record (VR). The address listings in the VR were used to control the delivery and return of questionnaires for all dwellings in the assigned area.

Respondents were asked to complete the questionnaire and return it by mail. Questionnaires returned by mail were, in turn, forwarded to the appropriate Census Representative (i.e. the Census Representative who delivered the questionnaire) for editing. The Census Representative checked each questionnaire for completeness. For each incomplete questionnaire, the Census Representative followed up by telephone and/or in person to obtain the necessary information. The Census Representative conducted non-response follow-up for each VR listing for which a questionnaire was not received through the mail within a specific time period. The Census Representative completed these questionnaires by personal interview.

Once the Census Representative completed enumeration for all listings in the VR and the assignment passed acceptable quality checks, all the questionnaires for that EA were forwarded to one of several Data Capture Centres. For the 2001 Census, data capture was contracted out to the Canada Customs and Revenue Agency (CCRA). CCRA has, in fact, been the data capture arm for the Census since 1981. This has been a very cost-effective solution since CCRA had idle equipment and trained available staff during their “slow” period after the tax filing deadline, which precisely coincides with the needs of the Census cycle. For the 2001 Census, manual coding of the industry and occupation variables was also contracted out to CCRA, once again for the reasons mentioned earlier.

Statistics Canada introduced imaging of the questionnaires in 2001 and all the subsequent steps of editing and correction were conducted using images rather than having to handle physical forms. For over 30 years now, our

[1] Anil Arora, Director, 2006 Census Manager, Statistics Canada, 1-B2, Jean Talon Bldg, Tunney’s Pasture, Ottawa, ON, K1A 0T6, arorani@statcan.ca

approach to conducting a Census has yielded high quality results. Robustness and cautious experimentation are essential qualities, given the enormity of the task (implicating every household in Canada, employing 45,000 staff, incurring considerable expenditure), and the fact that we only have one chance to "get it right".

Our approach of parcelling the country into manageable assignments, each fit for a Census Representative (Enumeration Areas), provides a sense of ownership, commitment, and pride in achieving good quality (coverage and completeness) at a fairly low geographic level. Conversely, if there was a problem, it could be contained to the same small area.

Until 2001, Statistics Canada consciously favoured a manual robust system for data collection over a "technology dependent" option, where certain aspects of the latter had yet to be proven in a Census context. "We have now reached the stage where a confluence of events pushes us in a new direction, just as technology reaches maturity, and where we believe the risks have diminished to a manageable level."[2]

2. THE DRIVERS FOR A CHANGE IN COLLECTION METHODOLOGY

2.1 Privacy

In 2001, as in previous Censuses, completed questionnaires were returned directly to the local census representatives for completeness checks and follow-up when necessary. Respondents are less and less comfortable with local census representatives having access to their personal Census data, especially so in rural areas. The Census is witnessing increased concerns expressed by Canadians through our Census help lines, calls to the Privacy Commissioner's Office and by our field staff. Privacy awareness is no longer limited to a few and our assurances of protection do not appease every Canadian household. Inaction in addressing privacy concerns would lead to a negative impact on data quality, response rates and potentially cost millions of dollars to conduct follow-up activities.

2.1 Internet option

With the growing use of the Internet for banking, commerce, and transactions with all levels of government, it is expected that by 2006, census respondents will demand an Internet option, and possibly other electronic options as alternatives to mailing back their paper questionnaire. Indeed, the Canadian government has stated its intentions to provide an Internet option to Canadians for all government services by 2005 (Government On-line).

2.2 Keying capacity

Census data have been manually captured by key entry at CCRA facilities since 1981. With the expansion of electronic filing of tax returns through E-file, Telefile and Netfile, we cannot expect the Canadian Customs and Revenue Agency to maintain its manual key entry capacity at a level and cost required to meet future Census requirements.

Intelligent Character Recognition (ICR) technology offers a proven alternative to manual key entry. The 2006 Census paper returns will be captured using optical scanning and image recognition (OMR and ICR) software. Capturing the responses as soon as individual questionnaires are mailed back, rather than waiting until an enumerator completes his/her assignment, supports the timely automation of the completion edits and computer-assisted telephone follow-up of failed-edit questionnaires. This approach, while providing prospects for reducing processing time, places it squarely in the critical path of collection activities and inter-twines two processes which traditionally were sequential and independent.

[2] Ivan Fellegi, Chief Statistician of Canada

3. OPPORTUNITIES

3.1 Large decentralized workforce

It is becoming increasingly difficult to recruit the large, temporary, decentralized workforce required to complete the census and to retain staff for the very short duration of the census in a competitive market place-offering higher remuneration and longer term employment prospects. The introduction of the new methodology will reduce our dependence on temporary employees. The 2001 Census employed over 45,000 field staff and the 2006 Census is estimating utilizing a field staff of 27,000.

3.2 Long-term cost control

A substantial portion of the cost of taking a Census is the field operations. Field staff is paid based on the number of dwellings and their wage reflects the salary and non-salary components of successfully enumerating each dwelling in their assigned area. The overall Census costs are thus very sensitive to factors such as the number of dwellings, wage rates, response rates, and economic increases in components such as fuel, transportation, and accommodation.

The 2006 Census methodology holds prospects for reducing the impact of such factors on the overall cost of taking a Census. As Internet participation rises, cost reductions should materialize. As data capture technology improves, less manual intervention will be required for data repair. As the Address Register coverage is extended, there can be increased mail-out to households, as opposed to census representative delivery.

4. DESCRIPTION OF PROPOSED 2006 CENSUS COLLECTION METHODOLOGY

The changes being proposed involve the re-engineering of the collection and data capture activities.

The major thrusts of the proposal are to:

- Develop a complete address file and to extend its coverage to areas including 80% of the housing units in Canada;
- Mail out in areas where considered feasible[3] (approximately two-thirds of all housing units in Canada);
- List/leave in non mail-out areas (approximately one-third of all housing units in Canada);
- Provide respondents options to return their data (via mail, Internet, telephone) directly to a Data Processing Centre (DPC);
- Introduce flow processing and an automated control register (processing and control by dwelling as opposed to census representative assignments);
- Automate registration, scanning and data capture of returns;
- Automate the editing of captured returns and for failed-edit returns, conduct computer-assisted telephone follow-up (CATI);
- Conduct field non-response follow-up (NRFU).

The focus of **pre-census day activities** will be coverage. This will involve updating the Address Register (AR) and delivering pre-identified questionnaire packages to all dwellings. Approximately eight months prior to Census Day, the AR will be provided to census representatives in order to conduct block canvassing. (Census representatives systematically canvass each block to locate every housing unit and compare each address with those listed in the AR.) They will make appropriate additions, corrections and deletions. This activity will last approximately eight weeks. An address file will be provided to the printers in order to prepare mail-out packages. In pre-identified mail-out areas, questionnaire packages will be mailed to householders approximately two weeks before Census Day.

[3] Areas would not be considered feasible for mail out where:
- addresses are not recognized by Canada Post (mainly rural and small urban areas);
- there is a great deal of instability and the Address Register is likely to be of poor quality;
- it would be operationally inconvenient, for example because there are relatively few mailable addresses surrounded by enumerator delivery areas (isolated pockets)

Statistics Canada will utilize updates from the geographic Street Network File, administrative records (building permits, Goods and Services Tax (GST) housing rebate files, telephone files) to identify any growth in these areas during the relatively short 6 month time period between block canvassing and mail-out.

In non-AR areas and in selected AR areas where mail-out is not feasible, Statistics Canada will continue to use census representatives to deliver questionnaire packages (list/leave) to achieve complete coverage. In these areas, census representatives will begin delivering questionnaire packages two weeks before Census Day. Respondents will have the option of returning their data via the Internet, through the mail, or by telephone. Respondent data will be returned directly to the DPC as opposed to field census representatives, therefore practically eliminating the local enumerator privacy concerns.

The focus of **post-census day activities** will be to ensure that a completed questionnaire is collected and processed for each dwelling. This will include capturing household data and following up on incomplete questionnaires and non-response cases. All questionnaires returned by mail will be forwarded directly to the DPC for automated registration, scanning/imaging, editing, and coding. All questionnaire data returned via Internet will automatically be registered and captured.

Data from scanned and electronic questionnaires will be subjected to a set of automated algorithms (completion edits) designed to verify if they pass the required edit rules. Cases that fail edit rules will be transferred to regional computer-assisted telephone centres for follow-up.

Non-response follow-up assignments (dwellings for which no questionnaire was returned by mail, telephone or Internet) will be generated from the control register two weeks after Census Day and activities managed from an estimated 33 Local Census Offices (LCOs) located throughout the country. Questionnaires completed by field staff will be logged at the LCOs and subsequently forwarded to the DPC on a regular basis.

4.1 Census Dress Rehearsal in May 2004.

As part of the risk mitigation strategy, a Dress Rehearsal is planned for May 2004. The major objectives of the Dress Rehearsal are to exercise the various systems and processes which are inter-dependant and critical to the success of the 2006 Census. As opposed to Census tests in previous Census cycles, the Dress Rehearsal is not an opportunity to "test", evaluate and select from a series of options. Systems and processes are planned, designed, developed, tested and implemented with the view to functioning as in 2006, differing only in the scale and scope of the Dress Rehearsal sample. In fact, systems are expected to be scaled up for 2006 and modified in design only in areas where major short-comings warrant changes. The Dress Rehearsal will also serve as the quantitative test for the 2006 Census questionnaire content.

Three concentrated areas of the country[4], comprising 300,000 dwellings have been selected for the Dress Rehearsal, mail-out and list-leave in roughly equal proportion. An additional 20,000 agricultural operations have been selected to ensure appropriate integration between the Census of Agriculture and the Census of Population processes.

Field operations with the block canvass activities (verifying and updating the Address Register in the field) begin this fall (September 2003), with the Dress Rehearsal Census day planned for May 11th, 2004.

[4] Atlantic Region - Halifax
Mail-out areas: parts of Halifax and Kentville
List-leave areas: parts of Annapolis Valley and Cumberland County
Quebec Region - Montreal or Longueuil
Mail-out areas: parts of Longueuil, Marieville and St-Mathias-Sur-Richelieu
List-leave areas: parts of the Montérégie
Prairie Region - Winnipeg
Mail-out areas: parts of Winnipeg, Selkirk and Yorkton (urban canvasser - part of Winnipeg)
List-leave areas: parts of central Manitoba and southeastern Saskatchewan

5. THREE KEY AREAS

Statistics Canada has taken an innovative approach to three key aspects of the 2006 Census: Content, the Internet as a data collection tool, and the Outsourcing of systems development and processing activities. The remainder of this paper deals with these three areas, including a status report of where we are with respect to each and highlighting some of the challenges and issues that lie ahead.

5.1 2006 Census Content

While the final content of the Census questionnaire is approved by Cabinet in Canada, questions submitted are based on extensive consultations with key stakeholders, with each change undergoing exhaustive cognitive, qualitative and quantitative testing. Some of the challenges associated with 2006 Census content are:

- willingness of the population to continue to provide sensitive data essential to fulfilling numerous legal statutes, regulations and data user needs
- managing the expectations of lobby and interest groups in placing questions on the Census, often for symbolic reasons, not specifically tied to legal or policy requirements
- retaining content from previous Censuses (time-series) while introducing new data needs and not contributing to respondent burden
- maintaining historical data comparability while conforming to emerging classification standards (e.g. SIC to NAICS, MFS to CIP, etc..)
- uncertainty in the funding necessary to collect, process and disseminate Census information
- maintaining content data quality while introducing major changes to the collection and processing methodology
- minimizing the potential impacts posed by collecting data through various modes (paper, internet, telephone)

Being cognizant of the risks associated with the magnitude of change in collection and processing methodology for 2006, Statistics Canada articulated a content consultation strategy based on a minimum-change approach. The basic elements of the strategy are:

- the provision of alternative data sources (surveys) which could satisfy user needs, sometimes with little or no modification to current content or sample sizes
- integrating consultations on 2001 data dissemination needs, geography requirements and 2006 Census content requirements
- adopting a longer horizon to anticipated data needs (post 2006 Census), and
- the establishment of a “business case” approach to internal (within STC) and external content addition or modification requests

This approach has proven successful in maintaining the relevance of the consultation process while minimising requests for content additions and modifications. The stakeholders consulted reported appreciating the integration of discussions of content, geography and output into a single efficient and effective process. They concurred with the context and were willing to explore other data sources to address data gaps where appropriate in lieu of introducing major content changes.

A short list of content changes and enhancements were recommended for consideration. The most broadly requested items were for:

- a new question on wage rates,
- updating the set of education questions (including a question on place where highest degree was obtained),
- a new question on duration of journey to work, and
- a new question on general health status.

In addition, two items for consideration were raised internally within Statistics Canada:

- providing respondents the option to consent to the use of tax filing information in lieu of completing the detailed income question
- seeking permission to allow access to census information after 92 years[5].

These six items have formed the core of the 2006 Census content testing program, complemented by a number of more minor changes to existing questions. The testing program consists of a cognitive testing component (focus groups and qualitative one-on-one interviews) and a quantitative component. The quantitative component will consist of a detailed analysis of the data collected during the 2004 Dress Rehearsal.

Of the items enumerated above which are being cognitively tested for the 2006 Census, the wage rate, general health status and the journey to work questions were dropped from further testing in the spring of this year. Early rounds of cognitive testing revealed potential data problems for both questions. It appears these questions would necessitate detailed, unacceptably lengthy instructions, and likely need more than one question to obtain meaningful results.

The remaining items (consent to use tax data for income, education module and permission to release) are still being tested. Should cognitive testing suggest that good quality data can be expected, they will be included on the Dress Rehearsal questionnaire. Final decisions on the inclusion of the questions into the cabinet submission document would follow the analysis of the data from the 2004 Dress Rehearsal.

In addition to the content determination process, the introduction of intelligent character recognition necessitates substantial format changes to the Census questionnaire. The 2001 Census long form (2B) did not segment write-in responses and requested information from 6 members of the household in a matrix rather than a sequential format on one questionnaire. The inclusion of segmented boxes for 2006, while increasing ICR scanning accuracy significantly, only leaves sufficient room on the questionnaire to accommodate 5 members[6]. This reduction in the number of household members accommodated by a single questionnaire, increases the number of households requiring an additional questionnaire (more than 5 members) to approximately 250,000 for 2006.

The 2001 Census short form (2A) is an English/French tumble format booklet (English on one side and French on the other). A number of issues associated with "bleed through" of responses, preparation of questionnaires for insertion into scanners, scanning millions of blank pages and associated disk storage requirements and associated costs, force the investigation of a single-sheet folded short-form. A number of qualitative tests are currently underway to investigate the possibility of migrating from a bilingual booklet to a unilingual single-sheet 2A format.

5.2 2006 Census on the Internet

Canada's Government On-Line (GOL) initiative states that by 2005, all transactions that Canadians make with the federal government should be available on-line. In complying with this initiative, recognizing the public's growing use of the Internet and meeting expectations for an on-line option, Statistics Canada (STC) will provide all Canadians the option to fill out their census questionnaire on-line in 2006. While in principle, this appears to be a relatively straight forward initiative, in applicability-poses some significant challenges in the areas of integration, communications, security, scalability, cost and design. The challenge is augmented by the fact that there is limited experience (in Canada and abroad) to draw upon in a full-scale Census context, other than small tests and contained offerings.

The 2001 Census in Canada included a small pilot study in two small geographical areas to test people's reaction to, and explore issues related to the provision of an Internet option. While the uptake was only about 5% of the eligible households in the two areas, Statistics Canada obtained valuable feedback on this process, guiding the development of the Internet option for the 2006 Census. The main lessons learned, and objectives of the 2006 Census application are:

[5] This is in response to an anticipated piece of legislation expected before the Canadian Parliament.

[6] A pre-test of formats (involving 21,000 households in April 2002) revealed that the matrix format was still preferable to a sequential format from a quality, timeliness and respondent friendliness perspective and met ICR requirements.

- the need to integrate systems and field processes to account for the multiple response modes and timely notification of responses
- the necessity of providing an application free of any downloads or permanent "footprints" on the respondent's machine
- the need for a simple and efficient application
- the use of standard web design elements in the questionnaire
- the provision for rapid scalability
- the need for secure data transmission to and from respondents
- the need for an effective public communications campaign promoting the "added value" to the respondent of using the internet option

By 2006, it is expected that between 60% and 70% of households will have an Internet connection and that between 30% and 40% of these households will choose to complete their census questionnaire on-line. Should these assumptions hold true, we can expect to receive approximately 18% of all households to respond using the Internet in 2006. Our budgeting assumptions are based on this projected take-up rate. It is important to note that more than 32% of households would have to respond over the internet in order to off-set the costs of developing and integrating the application. Given that investments have to be made in advance to handle the anticipated remainder of the responses on paper, a much higher uptake does not necessarily result in enormous savings.

Many Canadians already use the Internet to file their tax returns, to do their banking and to purchase goods on-line. In 2001, more than 5.8 million households, or 49% of all households, had at least one member that regularly used the Internet from home. This was up 1.1 million (+23%) from 2000. This followed an even stronger gain of 1.4 million (+42%) from 1999 to 2000 (source: Household Internet Use Survey, STC). More and more, Canadians expect an electronic option for most government services. The Canadian Customs and Revenue Agency (CCRA) received over 10 million (over 40%) tax filings in electronic form in 2003 (for tax year 2002)[7].

Canadians use the internet because it provides them some benefit-quicker tax refunds for CCRA, convenience and timelines for banking and purchasing goods-where there is a tangible benefit. The challenge for the Census is to find ways to provide similar incentives for the respondent.

Advances in security solutions will allow STC to provide a robust, user-friendly application with levels of protection for respondents' data that have been difficult to achieve in the past.

5.3 Application Development and Integration

In keeping with the census' strategy to contract out those activities which the private sector has demonstrated it can perform successfully, the development and implementation of the Internet option has been included in a contract awarded to Lockheed Martin. Lockheed Martin, the prime contractor, has chosen IBM to be the subcontractor in charge of establishing an application design, in partnership with STC. IBM will then develop and implement this design. By hiring an application integrator like Lockheed Martin, which will design and develop the majority of the processing applications, risks and issues related to integrating data collected from multiple channels will be minimized.

One of the key changes to the collection methodology that will allow the Internet to be more easily integrated is the use of an automated control register. By having a database with a placeholder for every dwelling before Census Day, we will be able to print a unique Internet access code on every questionnaire.[8] This access code will allow respondents to be authenticated, enabling them to use the census' data encryption software (which will provide a secure connection between the respondent's Internet browser and the census' servers) and to be presented with the appropriate form type, and will allow the census to register the household's response automatically. This will allow

[7] It is important however to note that CCRA offers a quicker tax refund where applicable for those who file electronically. The Census by contrast cannot provide such a tangible incentive for electronic reporting and thus it would be misleading to make comparisons of receipts using this method.

[8] Individual questionnaires and questionnaires for overseas households will not have preprinted Internet access codes. Respondents will be able to obtain an Internet access code from the Census Help Line.

us to eliminate some unnecessary follow-up, thereby reducing costs and respondent burden. The responses are edited while they are being captured, which will further reduce failed-edit follow-up. The electronic response will be automatically fed into the data-capture stream, and integrated with responses captured from paper questionnaires.

5.4 Cognitive Testing

STC's effort to design a user interface that would minimize mode effect, while providing the positive user experience necessary to achieve desired uptake levels, began with our small test of 200,000 dwellings in two areas during the 2001 Census. At that time, in order to achieve the desired levels of security for respondent data, Statistics Canada implemented a downloadable application that users installed on their computer. In addition to the lessons learned from that test, STC commissioned a series of tests, which were conducted by specialists from Carleton University's Centre for Human-Computer Interaction. Lockheed Martin/IBM's "Innovation Centre", which has successfully designed some of the world's largest and most critically acclaimed Web sites, is working with STC to design the user interface.

Another key component of our approach to designing the application's interface has been to integrate the good ideas of, and avoid the pitfalls experienced by, some of our colleagues around the world. STC has maintained a dialogue with representatives from statistical offices in Switzerland, New Zealand, Australia, the US, and Singapore, and has shared best practices and lessons learned with regard to collecting census data over the Internet. In order to promote this communication, Statistics Canada has organized a conference on Web-based Census Data Collection. The conference will be held in the Ottawa area in October 2003.

5.5 Security

Among the original drivers for considering changes to census collection methodology were the related issues of confidentiality (in particular, concerns over the use of local enumerators) and security of personal information. These are increasingly of concern to Canadians, whether it be with regard to public or private institutions. Respondent data entered via the Internet shall be protected through approved methods, including authentication and secure two-way (client-to-server and server-to-client) encryption of transactions. This involves leveraging GOL's "Secure Channel" – derived from the Public Key Infrastructure (PKI) concept – and the implementation of specific features, which will include the following: use of anonymous, limited-use certificates; isolating Statistics Canada's internal network through the use of an "air gap" concept; and ensuring that the application and the services that support it have the requisite capacity to handle the peak volumes of the census.

5.6 Application Features

In order to achieve its desired uptake targets, Statistics Canada will need to encourage respondents to use its Internet application. This will be accomplished by promoting the Internet option on the front cover of the questionnaire, as well as through STC's public communications and marketing campaigns.

The next challenge will be to ensure that as many of the respondents as possible who try the application are able to understand how to use it, and find it a pleasant enough experience that they finish their questionnaire. To that end, we have included a number of features to increase the application's usability. For example, there will be a simple, single-step logon process, during which the respondent will enter a code from the front of his or her paper questionnaire, and after which the respondent will be presented with the appropriate form type.

Respondents with "long" forms will be able to suspend their sessions and later resume where they left off. Respondents will also be able to resume their session from a different machine. For example, they could start their questionnaire at home, suspend their session, and then resume it from an Internet café. The application will have automated skip patterns and on-line edits, and will use respondent input to customize the questionnaire. For example, if the respondent indicates that there are three individuals living in the dwelling, the questionnaire will create three spaces for responses.

5.7 Outsourcing

With the need to find a cost effective replacement for manual key entry of Census questionnaires, and the maturity of Intelligent Character Recognition technology, the decision to seek external expertise was made early on in the planning for the 2006 Census. With the success experienced by other countries in using external expertise for data capture and processing Census returns (United States, United Kingdom, Australia), it was apparent that risk could be effectively transferred and managed by outsourcing. A set of outsourcing criteria were developed, ranging from cost, timeliness, risks, long-term impacts to quality and knowledge transfer-and applied to the range of systems development and processing activities required for 2006. A clear demarcation was made to restrict any contact with respondents to Statistics Canada employees, even though any contract staff would undergo all the necessary security checks and be sworn in under the Statistics Act-thus subject to the penalties under the law for the disclosure of any confidential information. By applying the criteria to individual and packages of inter-related activities, a set of activities to be outsourced was proposed to and accepted by Statistics Canada senior management. Concurrently, private industry was consulted heavily in their ability, expertise, and experience in the various activities being considered for outsourcing. The final list included:

- Printing of Census questionnaires in order to ensure that the type of paper, ink, folding and other characteristics met Intelligent Character Recognition (ICR) requirements
- Automated data capture of all paper questionnaires
- Coding of all write-in responses
- Operations of the Data Processing Centre, including facilities and human resource management[*]
- Development, deployment and maintenance of the Internet application as well as the Computer Assisted Telephone Interviewing (CATI) application.

The package of activities was appealing from a number of perspectives. Industry appreciated a diverse set of activities which were inter-related-where they could manage the interfaces and thus control integration risks. The package is also intellectually challenging, thus attracting the attention of industry leaders to form consortia with proven expertise.

The management of the activities with the external vendor is in a project environment with integrated project teams (vendor and Statistics Canada representation). The Outsourcing Project is divided into three distinct phases. Phase I comprises a detailed design phase allowing for a clear understanding of requirements and the presentation and documentation of a detailed solution outlining quality, risks, costs and schedules. Phase II comprises development, implementation and evaluation of the systems and processes for the Dress Rehearsal. Phase III includes the systems and process modification, scale-up and deployment necessary for the 2006 Census. A formal approval process (involving Government of Canada central agencies) is formally required before proceeding from one phase to the next.

Statistics Canada began industry consultations in 2001, formally released a Request for Proposal (RFP) in October 2002, and after a rigorous evaluation period, awarded the contract to Lockheed Martin as the prime contractor in February 2003. In June 2003, we received approval to proceed from Phase I to Phase II. Each phase has a well defined set of requirements to be met within a fixed cost ceiling, with incentives and disincentives based on specific cost, quality and timeliness milestones.

6. OUTSTANDING ISSUES

For each of the three areas highlighted in this document namely content, internet and outsourcing, a number of issues arise.

Several questions on the Census, although important and relevant to some specific groups, appear to have more of a symbolic significance rather than a direct policy, legal or program requirement. While in some instances, they serve as important catalysts in obtaining a relatively high interest and thus response rate, how should user expectations be

[*] This part of the contract is currently under consideration for the 2006 Census.

managed relative to policy/program requirements and response burden? In addition, in Canada and some other countries like Australia and New Zealand, a traditional Census is conducted every five years. While statistical systems and data users have wrapped needs around this historical frequency, central funding agencies exert pressures and uncertainties that add risk to such an enormous program. What measures can or should be taken to align the requirements of data users, with the fiscal constraints and requirements imposed by central agencies each Census cycle, while introducing more cost-efficient collection and processing measures?

Given the Government on-line initiative a number of issues arise from establishing an internet option. Providing a secure on-line option to every household in the country requires a significant investment in systems development, public communications, help-line support, and a secure infrastructure. Arguably, in 2006, even with an 18% response rate over the Internet, Statistics Canada will not recoup its investment made in meeting this public expectation of an on-line response channel. Hard investments have to be made in paper handling and processing systems for the projected remaining 82% of the responses. What measures can be taken in order to maximize the investments made in both electronic and paper-handling systems? Moreover, a significant component of the investment in an internet application deals with the hardware and software required to handle huge response volumes during a very short period (peak volumes expected to be on Census day). No reliable prediction models exist in mapping response patterns to infrastructure investments. What should the investment profile look like in building the ideal infrastructure versus putting in place a public communications strategy which attempts to smooth out a response profile?

Providing services on-line is a key objective of the Canadian government, including the development and provision of a secure infrastructure. While other government applications may encourage the re-use of a public key/certificate for multiple transactions, the infrequent nature of a Census along with confidentiality concerns prevent full utilization of such provisions. In the long-term, to what extent should statistical agencies continue to demand a clear differentiation versus capitalizing on government wide trends and investments?

The Canadian Customs Revenue Agency has realized enormous success in trying to get Canadians to use on-line applications and tools to file their tax returns. One of the primary incentives for an on-line completion is a quicker tax refund. For the 2002 tax season, over 10 million Canadians (40%) filed their taxes electronically. While the Census aims for a 20-25% response rate over the Internet in 2006, there appears to be the lack of clear articulated incentive for Canadians. What should the key communications messages be for 2006 in encouraging Canadians to complete their questionnaires on-line?

The provision of systems and solutions in-house has an associated implicit informality. While not entirely undisciplined or without fixed schedules for deliverables, short delays are more readily tolerated and changes more than likely accommodated. In outsourcing key systems development and processing activities, it is felt that there will be a significant erosion of any informality or ability to make changes late in the process-without incurring huge costs. Many statistical agencies have enjoyed relative success in engaging external partners in the conduct of a Census. While the capacity in the private sector appears to be increasing in meeting Census specific demands, capacity within statistical offices is decreasing or remains undeveloped. In the long-term, what should be the right balance of the extent and type of in-house expertise vs. that demanded from the private sector? Should statistical agencies work cooperatively in ensuring sustainability both in-house and externally?

The three areas discussed in depth in the foregoing represent only some of the challenges facing us as we implement such fundamental changes to a program as large and complex as the Census. Field procedures have to be re-designed, and the experience built up over the years in the Regional field organization are not necessarily directly relevant to the new environment. Changes in data capture technology, centralized follow-up of edit failures, autocoding of industry and occupation write-ins, will all have impacts on error patterns in the data, which need to be identified and understood. All of this reinforces the need for thorough testing and evaluation and the key role the 2004 Census Test will play in this regard…