

Effective Control Systems in a Multimode Census Data

Collection Methodology

David Dolson

Statistics Canada

100 Tunney's Pasture Driveway,

Ottawa, Ontario, Canada

E-mail: david.dolson@statcan.gc.ca

Introduction

For most countries census data collection and its control has been largely a decentralized paper based operation. Enumerators listed dwellings and conducted interviews or dropped off questionnaires. They tracked response and did follow-up using their lists for control. For some special situations telephone data collection has also been done. But over the last couple of decades some national statistical offices (NSOs) have developed address registers to facilitate mailout methodologies. Also there has been rapid improvement in the availability, security, acceptance and use of Internet technologies by both NSOs and citizens. These factors have facilitated the introduction of multimodal census data collection methods that better address preferred response modes of different population groups. It has also required the development of centralized control processes that can interact with a large decentralized data collection staff on a near real time basis.

The next section reviews the collection methodology of recent Canadian Censuses and discusses some of the lessons learned. This is followed by a section presenting an overview of the 2011 Census data collection methodology. The last section briefly describes the major strategies being taken for its control and management.

Data Collection in Recent Canadian Censuses and Pressures for Change

Up to 2001 data collection was conducted essentially as outlined in the introduction. In addition, Statistics Canada had developed an Address Register (AR) that could be used for dwelling coverage improvement by enumerators in some parts of the country. Each enumerator was to reduce non-response to no more than two percent and local supervisors known as crew leaders, used progress feedback from enumerators to manage their resources to achieve this objective.

Major changes were implemented in 2006. A new Master Control System (MCS) was developed to track and control the progress of every questionnaire through every step of data collection and processing. An option to respond via Internet was made available for all and each questionnaire was preprinted with a unique Internet access code. The quality of the AR had improved such that after a field validation a few months prior to Census Day, questionnaires with preprinted addresses could be mailed out to about 70% of dwellings. All of these addresses and codes linking them to their questionnaires were placed on the MCS. In the remainder of the country questionnaire delivery was done using a list/leave methodology similar to that in previous censuses. Enumerators had to ensure that each address listed was properly linked to the specific questionnaire dropped off. Upon completion of listing, enumerators provided information to be loaded on the MCS on how many and which questionnaires had been used. The actual list was retained by the enumerator for use during non-response follow-up. All responses whether by mail back, Internet or telephone to the Help Line, went to a central location where they were checked in against the MCS. Crew leaders initiated non-response follow-up by accessing the MCS to create lists of non-responding dwellings and forwarding these to enumerators. They then tracked progress using their lists and sent daily updates to their supervisors so that

the MCS could be updated. As responses continued to be received and checked in centrally, enumerators were notified of these receipts via their local offices so that follow-up effort could be ceased for the corresponding addresses. Edits for completeness and telephone follow-up were conducted at the central site.

While overall the strategy worked and a good result was achieved it was not without difficulty. The control mechanisms worked as intended but the information they provided was often not timely enough resulting in a waste of enumerator effort for questionnaires that had already been returned. There were two problems. First, non-response lists inevitably included addresses whose completed questionnaires had been mailed but had not yet been received and registered as such on the MCS. The time lag for this could be up to ten days depending on the sender's locale. Second, once non-response follow-up had started notifications of receipts were faxed to crew leaders for forwarding to enumerators so that follow-up for the address in question could be ceased. The fax process did not always work well, causing crew leaders difficulty informing enumerators in a timely way. Another result was that supervisors and managers did not always have a good picture of the status of collection operations and so were unable to deploy available resources to best meet the two percent maximum non-response target, particularly in the late stages of follow-up.

So, building on the lessons learned from these observations as well as the results of other evaluations and testing further important changes have been made for 2011.

The 2011 Canadian Census Data Collection Methodology

Strategically, multiple modes of data collection will be used in an organized and well controlled way so as to obtain efficiencies and to target different segments of the population. This is to be done with reduced resources – particularly field enumerators – while maintaining quality and timeliness. Several important changes are being made to achieve these goals. The Address Register has continued to improve in quality, and so the mailout methodology will be expanded to about 80% of addresses. Building on the success of Internet data collection in 2006, we are targeting to double Internet use to 35% of responses in 2011. Supported by a strong communications program, non-respondents will be contacted in a set of carefully coordinated waves of reminders prior to the start of field follow-up. The objective is to minimize the workload and cost of follow-up and hence its need for enumerators. Improved procedures for dynamic management of non-response targets will be implemented. Finally, the registration procedures, the control mechanisms and the tools for communication amongst collection staff are all undergoing important improvements.

The data collection is split into two mailout groups and one list/leave group. Group 1, about 60% of addresses, includes those places with high Internet connectivity and a demographic make up most likely to respond to the Census via Internet. These addresses will be delivered an individual Internet access code and letter encouraging reply via Internet. For those not wishing to do so, instructions are included for contacting an automated service to have a questionnaire delivered to their address within five days. Group 2, about 20% of addresses, will be mailed questionnaires each with a preprinted address and an Internet access code. Again response via Internet will be encouraged. For both groups the mailout is to be received by respondents seven days prior to Census Day. Group 3 includes the remaining 20% of addresses; questionnaires will be dropped off using the same list/leave methodology as described above for 2006. Each questionnaire will have a unique Internet access code. This process will be done over the 7 day period prior to Census Day.

In a second wave two days after Census Day, non-respondents in Groups 1 and 2 will be sent individual reminder letters. At this time Group 3 addresses are available only on paper lists; all dwellings in group 3 areas will be sent a reminder card. In a third wave of contact starting about 10 days after Census Day, remaining Group 1 non-respondents will be mailed a questionnaire. There is no wave 3 contact planned for

Groups 2 and 3 but in an experiment in the 2009 Census Test, half of Group 2 non-respondents will be sent a second questionnaire while the other half will receive a voice broadcast reminder message.

The fourth wave is field non-response follow-up by enumerators. For Group 3 this will start 10 days after Census Day. For the other two groups it will start another 10 days later to provide time for response to the wave 3 contact.

Effective Control of Data Collection for 2011

As in 2006, a Master Control System will play a central role in providing information on the status of every questionnaire through all steps of questionnaire delivery, return, follow-up and processing. In addition, important improvements are being made via a more timely registration procedure, a new Field Management System (FMS), generally similar to that used by the Australian Bureau of Statistics for its 2006 Census, and a new approach for dynamic management of resources to meet non-response rate targets.

In a first major initiative to address lessons learned from 2006 and to help facilitate the above data collection methodology, Statistics Canada has partnered with Canada Post for logistics and mail delivery needs. Canada Post expertise will be important in the preparation and delivery of large quantities of mail over a very short period of time in each of waves one to three. More importantly for control of data collection, registration of receipt of paper questionnaires will take place at a set of 18 regional postal plants. Registrations will be cumulated and transmitted hourly to the MCS. This will provide for a major improvement in timeliness of questionnaire receipt registration over the 2006 Census experience.

As in 2006, registration of questionnaires completed via Internet and Help Line will be essentially instantaneous. Wave 4 field-completed questionnaires will be registered daily by enumerators using the FMS.

The Field Management System is an important innovation for Statistics Canada that will address the communication issues experienced in 2006 as well as providing several other features to support data collection. It is designed as a web based application accessible from office or home by authorized staff including enumerators, crew leaders, support staff and managers both in local offices and at the central office. Although no respondent data will ever be available or transmitted via the FMS, security remains very important as confidential paradata and other information will be accessible. Authentication of authorized users will be via a user name and password as well as correct response to a challenge requiring reference to a unique card provided to each user. Access to the functions and information available via the system will be permissions based. This will be managed via a Field Assignment Structure to which the FMS refers. This will indicate each employee's position, responsibilities (e.g. activities, geographic units) and what permissions each has regarding the various FMS functions and other systems with which the FMS interfaces. With this the FMS will be able to direct information where it needs to go -- from assignment lists, messages or notifications for specific employees or units of geography to broadcast messages for all collection staff. As well, an employee will be presented only with the information relevant to their current position and responsibilities. Supervisors will be able to adjust these permissions as required to adapt to operational needs.

Major FMS functions include: data capture of Group 3 address lists created by enumerators, creation and management of non-response follow-up assignments; access for enumerators to twice daily updates of assignment lists (notifications of questionnaire registrations); a facility to send action requests to enumerators and their supervisors; messaging functions; access to much more timely and precise management information than ever before on progress, costs, enumerator performance, etc.; access to a Wiki of enumeration directives and procedures; access to a shipments tracking system; and a facility for daily entry, review and approval of collection staff pay claims. Some of these are now outlined.

Using the FMS and its near real-time access to the information on the MCS, crew leaders will create and manage non-response follow-up assignments for wave 4 of data collection. Assignments can be created manually or automatically with user specified parameters and can then be assigned by the crew leader to their enumerators to access remotely. Daily, using the FMS, enumerators will be required to enter registrations of completed questionnaires (for updating of the MCS) as well as contact attempts at other addresses. As progress is made, assignment lists can be closed at any time and new lists created. Using FMS quick reference tools, crew leaders will be able to easily monitor progress by assignment, enumerator or geographic area. This will provide the essential information for crew leaders and managers to monitor progress and performance.

Updates on questionnaire receipts, whether registered at a Canada Post site or at the central office, will be sent twice daily from the MCS to the FMS. Since the FMS keeps track of which addresses have been assigned to which enumerators, notifications of these receipts will be available to the correct enumerators the next time they access the FMS. This ease of use and timeliness will be an important improvement for 2011.

The FMS will provide a powerful, flexible, and easy to use capability for crew leaders, operations supervisors and other managers to access precise and timely management information. Dashboard type indicators will provide a quick means to examine expenditures and the state of collection progress for their areas of responsibility. These indicators will also be available in the form of reports with capability to drill down to low levels of detail. Staff summaries will be available to crew leaders to facilitate tracking assignments, notifications, action requests, pay claims, etc.

Finally, the FMS will play an important role in a more active and responsive management of collection resources to meet non-response rate tolerance limits than ever before. In past Canadian censuses this tolerance management took place almost entirely at the local level with a uniform limit of two percent non-response and near uniform staff levels and ratios in every enumeration area. Up to 2001 this approach worked well. In 2006, with a reduced mail return rate this became much more difficult and the final non-response rate was higher with a larger difference between the best and worst local areas even though improved management information was available through the 2006 MCS. For 2011 the initial two percent objective will remain the same but staff will be allocated differentially – using information from the 2006 MCS - to account for expected collection difficulty. The FMS will include a tolerance management system that will provide information on progress towards non-response tolerance limits. It will incorporate models for forecasting non-response follow-up progress as a function of collection difficulty indicators, initial response rates, local staffing levels, rates of collection progress, time remaining in the schedule, etc. It will automatically send indicators of when tolerance has been met in each area so that local collection managers can redirect resources where they are most needed. As well, using information from the FMS to not only monitor progress in general but also to identify local areas where progress is much better or worse than average, a non-response management team at head office will be well positioned to best meet the objective of minimum overall non-response with better more balanced control of non-response at local levels by directing re-allocation of field resources and adjusting non-response tolerance limits as needed.

Conclusion

For its 2011 Census, Statistics Canada will extend its use of a multimode data collection methodology. A structured series of waves of contacts with respondents is designed to maximize response via Internet and response in general prior to starting non-response follow-up. The Master Control System and Field Management System create a flexible, powerful approach to provide timely information and communications for dynamic and adaptive control of widely dispersed data collection operations and staff.