

**UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE**

CONFERENCE OF EUROPEAN STATISTICIANS

Seminar on New Frontiers for Statistical Data Collection
(Geneva, Switzerland, 31 October-2 November 2012)

Topic (v): Economies of scale from using common tools and methods

**ELECTRONIC QUESTIONNAIRE COLLECTION AT STATISTICS
CANADA**

Contributed Paper

Prepared by Milana Karaganis and Marc St-Denis, Statistics

I. Introduction

1. At Statistics Canada, electronic questionnaire (web questionnaire or e-questionnaire (EQ)) has been long recognized as a very important reporting option for all respondents and surveys. Various electronic data reporting tools have been explored over the last 20 years (Mayda (2002), Burgess (2007)). Statistics Canada has continued its attempts to develop an effective and efficient e-questionnaire solution that will meet respondent expectations, comply with Statistics Canada requirements for confidentiality and security, and comply with Government of Canada requirements for accessibility and common look and feel.
2. In 2010, a new corporate project was put in place to implement e-questionnaire as the primary mode of collection for over 160 business and household surveys. An integral part of this project is the use of a corporate e-questionnaire solution that will allow a standard approach to e-questionnaire development and data collection. Use of a standard corporate solution (as a common tool and method) is also expected to bring efficiencies and further reduce costs of data collection.
3. This paper will provide an overview of key aspects and features of the implemented standard e-questionnaire solution (tools, methods, processes, e-questionnaire design standards, collection strategies) and present a project status update, including realized efficiencies.

II. E-Questionnaire Project

A. Project objectives

4. The e-questionnaire project was centred on the new approach to electronic data reporting (Faid (2009)) with the use of web-based questionnaires. This approach was formalized in 2010 when a corporate e-questionnaire service was launched for use by all Statistics Canada surveys. E-questionnaire was targeted to be implemented as the primary collection mode for the majority of Statistics Canada surveys. Correspondingly, survey questionnaires would have to be designed to make e-questionnaire the primary mode of collection and the first step in a sequential multi-mode collection approach.

5. The following project objectives were established:
 - (a) To create a critical mass of e-questionnaire service users by deploying e-questionnaire for 160 business and household surveys;
 - (b) To establish standard and robust workflows for the e-questionnaire collection mode;
 - (c) To establish design standards and a standard approach to design of e-questionnaire applications;
 - (d) To upgrade functionality available for e-questionnaires in response to survey needs (as reviewed, prioritized, and approved based on Statistics Canada corporate needs);
 - (e) To establish collection strategies for multi-mode collection involving e-questionnaire collection mode;
 - (f) To further reduce collection costs;
 - (g) To work towards a model where the e-questionnaire system is used for interviewer assisted collection as well.
6. The plan for the 2016 Census e-collection is to migrate to and use this new corporate e-collection platform.
7. The main drivers for this project are:
 - (a) *Response rates pressures and respondent expectations* – by offering a collection mode that is convenient and popular with respondents (and is expected to be offered), it is believed that we could further improve respondent cooperation and reach population sub-groups that are more difficult to contact and gain cooperation with more traditional collection modes (e.g., young males).
 - (b) *Costs containment* – Internet take-up rates of 50% for business surveys and 20% for social surveys are targeted, which should lead to reduction in costs associated with interviewer-administered surveys.
 - (c) *Reduction in development time* – by using a standard e-questionnaire platform, we will have a faster turnaround period for development of an e-questionnaire application when compared with a Blaise application development.
 - (d) *Data quality* – use of soft edits, automated skip patterns, presenting definitions of key concepts and other immediate help functions are believed to ensure that respondents understand survey questions and provide quality answers.

B. E-questionnaire platform

8. The project makes use of an e-questionnaire platform that is secure; meets the requirements of the Government of Canada with respect to accessibility and Common Look and Feel (CLF); transparent for respondents (zero footprint on a respondent computer); respects usability principles; and is open, flexible and scalable from a technical standpoint.
9. The e-questionnaire platform was developed in-house and includes three large components:
 - (a) *Electronic Questionnaire Generation System (EQGS)* – EQGS is a custom-built system that allows e-questionnaires to be automatically generated from a generic dynamic specification template, which includes questions, flows, question help text as well as a choice of controls to be used by different questions (e.g., drop-down menu, radio buttons etc.). EQGS allows reduction

in development time by eliminating manual programming of e-questionnaire applications. Implementation of EQGS required standardization of functionalities offered for e-questionnaires for all surveys, and consequently, led to a reduction in customization of collection tools built for different surveys. This is a significant change for survey managers as surveys are now expected to create survey questionnaires that use generic functionalities within EQGS, rather than developing one-off solutions.

- (b) *E-questionnaire systems* - once an e-questionnaire application is created, it is placed in the e-questionnaire systems that support an overall user experience and manage e-questionnaire collection. The e-questionnaire systems include the *E-collection portal* (a landing portal for respondents attempting to access an e-questionnaire); the *Survey Management System* (survey initiation, sample, etc.); the *Data Management System* (data transformations that are required for re-integration of e-questionnaire data with other Statistics Canada systems); the *Management Information System* (reports on e-questionnaire data collection activities); etc.
- (c) *E-questionnaire infrastructure* - includes the hardware and software required to support the e-questionnaire platform.

C. E-questionnaire business processes

10. Standardized and generic business processes to support the e-questionnaire mode are key to delivering expected cost savings and include the following:
 - (a) *E-questionnaire standards* - EQGS translates supplied inputs into a final e-questionnaire application based on established functionality and look-and-feel standards. Thus, there has to be a process to establish a set of e-questionnaire standards that have to be maintained and updated based on newly identified survey needs.
 - (b) *Change management process* - a process is required to govern assessment of new functionality requirements, priority setting, and content determination of new releases of all components of the e-questionnaire platform. To respond to this requirement, a change management process was put in place to react to all requested changes (including not only e-questionnaire platform functionality change requests but also change requests regarding adjustments to business processes, e-questionnaire deployment scope for a particular survey, etc.).
 - (c) *Consultation with survey managers* - since the e-questionnaire platform is a generic corporate system that will be used by all surveys with the e-questionnaire collection mode, a forum was established to enable communication and collaboration between the e-questionnaire project team and the entire subject-matter community (as opposed to one survey at a time).
 - (d) *Standard collection strategies* – standard collection strategies are required to establish proper multi-mode collection procedures and balance collection costs against quality of collected data and resulting estimates. Collection strategies determine which collection modes will be used, sequence of collection modes, cut off points for a switch between collection modes, as well as non-response and failed-edit follow up strategies (i.e., number of attempts, attempt modes, timing of follow up attempts).
 - (e) *Standard e-questionnaire deployment process for surveys* - the main objectives were to reduce the deployment cycle and to create a standard set of steps to be followed by all service providers. The deployment process includes such steps as initialization (determination of e-questionnaire strategy – how it will fit with other collection modes, what standard functionalities will be used by an e-questionnaire application, how the entire collection process will work, etc.); building new e-questionnaire functionalities if required and approved through the change management process; designing and rendering an e-questionnaire application; mapping an e-questionnaire application to a Blaise application for CAI / paper collection; testing; establishing collection strategies; training e-questionnaire Help Desk and interviewers; preparation of respondent materials; loading of EQ sample files; preparation and mail-out of e-invitation and e-reminders.

III. Project Status

A. Progress update

11. The e-questionnaire project was initiated in spring 2010. The new e-questionnaire platform became operational in March 2011, and the first nine business surveys started collection in the new platform in April 2011. Since then, 30 more surveys started e-collection in the new platform, for the total of 39 surveys. This includes business surveys, institutional surveys, agriculture surveys, as well as the first household survey that used e-questionnaire as a collection mode (General Social Survey – Caregiving cycle, Internet pilot). While many of these surveys continued to use e-questionnaire in combination with other collection modes (multi-mode collection), others opted to use e-questionnaire as the only collection mode.
12. In parallel with e-questionnaire deployments, new multi-mode collection strategies for business surveys were tested through an embedded experiment (Karaganis, Fox, Claveau, Leung and Lin (2011)). This led to development of standard follow-up strategies that will be further discussed in Section IV of this paper.
13. On the business process side, a set of e-questionnaire development business processes was created. Collection partners were actively involved in the process development which helped with the training aspects as well. The fine-tuning of established processes and procedures will continue to allow for further improvement and streamlining of all business processes.
14. Implementation of the e-questionnaire collection mode coupled with implementation of new multi-mode collection strategies and processes resulted in generation of efficiencies in data collection. The e-questionnaire project is expected to generate on-going annual efficiencies of approximately \$2,400,000. After the first two years and deployment of the first 39 surveys (out of 160 surveys on the deployment list), the project already returned on-going annual efficiencies of close to \$800,000 and is on track to reach overall targeted efficiencies. The project is well on track to meet deployment dates for the remaining surveys with the target end date of spring 2016.

B. Use of e-questionnaire by surveys

15. Table 1 below summarizes key collection information for selected surveys that use e-questionnaire collection mode.

Table 1: E-questionnaire Survey Collection

	Survey type	Collection frequency	EQ sample	EQ submission rate for last cycle
Survey of Suppliers of Business Financing	Business	Two times / year	121 units (100% EQ)	97%
Sawmills	Business	Monthly	175 units (85% of sample)	70%
Unified Enterprise Survey (UES) – 7 industries	Business	Annual	6,850 units (73% of sample)	68% weighted response rate = 83%
Research and Development in Canadian Industry Survey (RDCI)	Business	Annual	1,473 units (75% of sample)	63%
Survey of Regulatory Costs Compliance	Business	Occasional	19,678 (100% EQ)	55%
Major Special Crops	Agriculture	Three times / year	30 units (88% of sample)	56%
Corn and soybeans	Agriculture	Three times / year	41 units (41% of sample)	63%
Public Service Employee Survey	Institutional	Once every 3 years	296,755 (100% EQ)	67%
Survey of Staffing	Institutional	Annual	87,257 (100% EQ)	48%
General Social Survey – Caregiving cycle; Internet pilot	Household	Annual	3,487 (61.7% EQ acceptance rate)	38.7%

16. Key definitions:

- (a) EQ sample column – EQ sample provided in terms of units as well as a proportion of the overall multi-mode sample. The only exception is the General Social Survey (GSS) where EQ acceptance rate is provided, which is defined as (units that accepted to complete questionnaire via EQ) / (total units contacted for the survey).
- (b) EQ submission rate is defined as a number of units that submitted e-questionnaires. For the UES surveys, weighted response rate is also indicated.
17. Business surveys were the main focus of this first wave of e-questionnaire deployments. Annual and quarterly business surveys as well as one monthly business survey started using e-questionnaire as the primary collection mode. Overall, Internet take up rates for business surveys were in line with expectations and some business surveys successfully used e-questionnaire as the sole collection mode. The next two major milestones for business surveys in the context of the e-questionnaire project will be November 2012 (beginning e-collection for the first mission-critical monthly survey, Business Payrolls Survey) and spring 2014 (beginning e-collection for the entire Unified Enterprise Survey program with more than 50 different surveys). Business surveys in the e-questionnaire project will be discussed in more details in Section IV of this paper.
18. For household surveys, the work has begun to deploy e-questionnaire collection mode for key household surveys (such as Labour Force Survey (LFS), General Social Survey (GSS), Canadian

Community Health Survey (CCHS)). Household surveys in Statistics Canada rely heavily on customized Blaise collection tools that are created for interviewers to complete data collection. Given the complexities of Blaise programming and the fairly complicated logic within each collection tool, typical household surveys do not have a paper version and a typical interviewer training session consists of two to three full days focused on survey concepts. Within the context of the e-questionnaire project, the complex nature of household surveys geared towards interviewer data collection represents an additional challenge as an extra step of translating the survey from an interviewer-completed instrument to a self-completed instrument is required.

19. Another important consideration for household surveys is sampling frames and methodologies and their impact on possible contact strategies that will aim at achieving respondent cooperation in completion of e-questionnaires. For instance, some household surveys use Random Digit Dialing frames with samples consisting of telephone numbers. In these circumstances, there has to be an initial collection phase to contact a selected sample unit, collect basic household information, select a respondent for the main survey, and secure respondent cooperation. Another example of a current methodological approach for household surveys is respondent selection within a sampled household. Many cross-sectional household surveys select a targeted respondent from the household roster that has to be first created by an interviewer in the first phase of an interview. Transitioning this approach into the Internet world, while allowing households to roster themselves, requires additional analysis to determine the implications for data quality, respondent compliance with collection procedures, and potential Internet take-up rate.

20. The first household survey to use e-questionnaire for data collection was piloted in summer 2012. The General Social Survey – Caregiving cycle conducted an Internet pilot in May-July 2012. This is a cross-sectional volunteer household survey on respondent’s caregiving experiences with a 45-minute subject-matter questionnaire. It uses a Random Digit Dialing sampling frame, so the data collection process was split into two phases. Phase 1 consisted of interviewers contacting selected phone numbers, establishing a household roster, selecting a respondent for the main survey and offering an e-questionnaire as a collection mode. For respondents that agreed to use e-questionnaire, an email address was collected and an invitation was emailed. In phase 2, respondents completed e-questionnaire and there was a non-response follow-up in the form of email reminders for outstanding units. After three email reminders, if respondents still did not complete e-questionnaire, these units were finalized as non-response. If respondents refused to use the e-questionnaire during Phase 1, the unit was finalized as a refusal. There were no telephone interviews done for the main survey. The table below summarizes key collection results for this pilot:

Table 2: General Social Survey – Caregiving cycle, Internet pilot

Sample (1)	CATI						EQ		
	Voids (2)	Net Sample (3) = (1)-(2)	Respo ndents to CATI (4)	Resp. Rate to CATI (4) / (3)	Accept to complete EQ (5)	EQ Acceptance rate (5) / (4)	Submit EQ (6)	EQ Submissi on rate (6) / (5)	EQ Respon se rate (6) / (4)
10,006	1,458	8,548	5,654	66.10%	3,487	61.70%	1,349	38.70%	23.90%

21. The first Internet pilot for a household survey was very successful in testing multi-mode collection mechanisms and delivering take-up and response rates above expectations. Collected data is being analyzed and assessed against CAI collection results from the collection cycle that was run in parallel with the Internet pilot. Meanwhile, work is under way to launch e-questionnaire collection for the next cycle of the General Social Survey, starting in the spring 2013.

22. In parallel, work is under way to launch an Internet pilot for the Canadian Labour Force survey (Karaganis and Laflamme (2012)) in January 2013. Labour Force Survey (LFS) is a mandatory monthly dwelling-based household survey with a collection period of 10 days each month where sampled households are interviewed for six consecutive months. Given differences in sampling and collection methodologies, LFS will use a different multi-mode collection strategy than the General Social Survey. The LFS sample design will remain the same while a new collection mode will be added. The first month will still be conducted via CAI and at the end of the birth interview, respondents eligible for e-questionnaire collection in subsequent months will be made an offer to complete the survey in subsequent months via e-questionnaire. Respondents who cannot do a survey over the Internet or refuse the e-questionnaire will remain in CAI. At the beginning of the next collection period, an invitation by e-mail will be sent to e-questionnaire respondents with instructions and an access code. The respondents will have three days to complete their e-questionnaire. At the end of these three days, e-questionnaire non-respondents will be sent to CAI for non-response follow-up. This tight time frame is required to respect the current release schedule while maintaining data quality.

C. E-questionnaire platform, functionalities and standards

23. As mentioned above, the e-questionnaire platform allows automatic generation of e-questionnaire applications based on specified inputs. This is accomplished with the help of the e-questionnaire generation system (EQGS) that reads provided inputs (specifications) and renders them into actual e-questionnaire applications.
24. To function properly, EQGS requires implementation of standard functionalities and e-questionnaire (EQ) design standards. The first version of our EQ standards was drafted in May 2011 and has been updated as new functionalities and standards were put in place. EQ standards describe design features of e-questionnaire applications – screen layouts, look and feel, visual display of different elements and controls. EQ standards do not cover traditional aspects of questionnaire design that pertain to content wording and ordering.
25. EQ standards have been developed by a multi-disciplinary team that includes questionnaire design researchers, business and household survey collection experts, system developers, and usability and accessibility experts. EQ standards have been developed to be fully compliant with existing accessibility and Common Look and Feel (CLF) requirements established by the Government of Canada.
26. With e-questionnaire applications being implemented for more and more surveys, standard functionalities available through EQGS have been upgraded in response to business needs. Currently, EQGS offers a wide spectrum of different functionalities to be used in e-questionnaire application designs. Examples of these functionalities include: standard pages to introduce a survey and provide instructions for completion; general help functions at the survey level and question-specific help functions; roster of surveyed entities (e.g., household members, business locations etc.); component list (to allow completion of multiple components within a questionnaire and facilitate navigation between components); progress bar to indicate survey completion progress; Stop & finish later function; calendar widget and dates answer type that allows mathematical calculations on selected dates; auto-sum within a column; two-column tables; edit checks and validation messages; answer type controls - grids; radio buttons; check boxes; drop-down menus; text fields; comment boxes; questions with sub-headings; different answer types (numeric; dates; currency; non-negative currency; integer etc.); and many others.

IV. E-questionnaires Among Business Surveys

A. Review of EQ Collection Strategies

27. This section will focus on the implementation of the EQ among business surveys at Statistics Canada, how the EQ efficiencies are being achieved, and some collections results from these surveys. Currently the majority of business surveys use paper questionnaires that are mailed out

for data collection. In some cases, a telephone pre-contact is done to verify contact information and confirm if the business is in scope. Data capture is done from mailed back questionnaires. Failed edit follow-up (FEFU) is conducted to resolve edit failures from mailed back questionnaires. Also, telephone non-response follow-up (NRFU) is done with respondents who have not returned the questionnaires in a pre-determined time period. As described in a previous section, managers overseeing collection were tasked with developing an EQ service with the goal of moving towards EQ collection as the primary mode for all business surveys over the next 5 years.

28. The first business surveys introduced to EQ have demonstrated the potential for electronic data reporting to supplant more traditional collection modes and achieve cost savings. In the initial surveys, the EQ take-up rate was limited, although promising. In reviewing the collection strategies for these initial surveys into EQ, it was noted that opportunities existed to improve efficiency and reduce costs. Towards the expansion of multi-modal collection and ensuring cost efficiencies targets are achieved, a review of the results of surveys using EQ was conducted in the latter part of 2011 for half dozen surveys. The results have shown some successes and identified areas requiring improvements. From the review, recommendations for optimum EQ collection strategies were made while meeting the response rate and quality targets, considering respondent burden, and collection timeliness. Part of the challenge was to convince the survey clients in adopting a collection strategy geared to EQ. It was observed that although some survey managers welcome the EQ collection mode, they prefer integrating EQ in the existing survey collection strategy rather than adapting a strategy specifically for EQ. In one case, the client managers were insistent on including in the e-questionnaire a notice informing the respondent that they could mail or fax their financial statement rather than completing the financial questions. This resulted in receiving over 1700 financial statements (35% of the sample) by mail or fax. For another business survey, the client insisted in conducting a courtesy call verifying that the respondent had indeed received the e-questionnaire. Also the collection strategy included the mailing of a post-card reminder as a second collection effort to all non-response cases rather than sending an e-reminder.
29. The Quarterly Services Indicators Survey (QSI) is one of the first surveys in Statistics Canada introduced to EQ multi-modal collection. EQ results are shown for the initial survey collection approach, and then after it was modified based on recommendations. The EQ uptake was very high for this survey, being at 98%. Table 3 provides an overview of the QSI collection strategy for quarter 1 of 2011. Telephone follow-up was conducted on all cases including the EQ cases before sending any e-reminders. A fax reminder was sent to NRFU cases. Once again the review conducted in the December of 2011 shown that this approach was sub-optimal.

Table3: Quarterly Service Indicators Survey, collection strategy 2011 (EQ+CATI)

Mail out / E-invitation	Oct 3, 2011
1 st e-reminder / fax	Oct 17
2 nd Phone reminder (all)	Oct 24
3 rd Phone reminder (all)	Nov 7
4 th e-reminder / fax	Nov 14
5 th e-reminder / fax	Nov 28
Fax to EQ non-response	Dec 5
End of collection	Dec 9, 2011

Table 4: Quarterly Service Indicators Survey, collection results Q1 2011

	EQ		Mail		Total
Final Sample	1,525	98.3%	27	1.7%	1,552
Initial Sample	1,525	98.3%	27	1.7%	1,552
Response Rate	58.5%		50.0%		58.0%

FEFU (hours)	41.2	97.4%	1.1	2.6%	42.3
# cases in FEFU	502	32.9%	13	48.1%	515
Avg min per FEFU case	4.9		5.1		4.9
NRFU (hours)	125.3	97.9%	2.7	2.1%	128.0
# of cases in NRFU	905	59.3%	22	81.5%	927
Avg min per NRFU case	8.3		7.4		8.3
Bounce backs	129*	* auto-reply bounce-backs were excluded			

30. Following a review of the QSI collection strategy, changes were brought to maximize the EQ advantages: e-reminders were made for the 1st, 2nd and 3rd attempts, followed by a phone follow-up as the 4th attempt. A 5th e-reminder was sent followed again by telephone attempts. Table 4 shows the metrics for QSI quarter 1 2011 and table 5 shows the metrics for QSI quarter 1 2012. Comparing the results between Q1 2011 to Q2 2012, the number of cases completed by EQ increased by 8%; also the number of cases requiring NRFU was reduced from 59.3% to 50.6%. It can be deduced that the changes made to the collection approach were successful in reducing the NRFU attempts thus yielding cost efficiencies.

Table 5: Quarterly Service Indicators Survey, collection results, Q1 2012

	EQ		Mail		Total
Final Sample	1,351	96.4%	50	3.6%	1,401
Initial Sample	1,364	97.4%	37	2.6%	1,401
Response Rate	66.4%		46.0%		65.7%
FEFU (hours)	44	91.1%	4.3	8.9%	48.3
# cases in FEFU	414	30.6%	24	48.0%	438
Avg min per FEFU case	6.4		10.8		6.6
NRFU (hours)	94	93.6%	6.4	6.4%	100.4
# of cases in NRFU	683	50.6%	41	82.0%	724
Avg min per NRFU case	8.3		9.4		8.3
Bounce backs	113*	* auto-reply bounce-backs were excluded			

B. EQ Collection Strategy Guidelines

31. In an effort to ensure more optimal collection strategies for surveys using EQ, the a Collection Strategy Guidelines (Appendix 1) were established to help collection managers prepare a follow up schedule that would facilitate collection across all collection modes, reduce respondent burden and in turn, would generate efficiencies. When deciding on the collection strategy to be utilized for a survey, there are a number of elements that must be considered: the mode(s) of collection being used; the collection period; the subject and content being collected; the type of follow up to be utilized; allocated budget; respondent burden; response targets; etc. These elements were included in the guidelines template and can be utilized as needed. The goal of these guidelines was to present a more standardized approach to our collection process and thus harvest efficiency in the collection process. A guideline table was prepared for different collection periods: annual, semi-annually, quarterly, and monthly. The table below is a guideline to be used for annual surveys. Over time, the results from this strategy will need to be evaluated to determine if the efficiencies and quality targets envisioned are realized.
32. As previously mentioned, Web based surveys were introduced for a number of reasons: a government-wide initiative to advance the use of the internet and its associated technology; increased requests from our survey respondents to provide them with an electronic self-reporting option; a vision to provide a more environmentally friendly alternative to paper surveys; and an

opportunity to achieve efficiencies in the collection process. Efficiency targets are set for each survey prior to EQ introduction. The efficiency targets have been determined as being approximately 25% to 30% of the interviewing costs. Collection managers are required to harvest these efficiencies. These cost savings are derived from different collections processes: questionnaire printing; postage; capture costs; CATI efforts for NRFU and FEFU. The metrics (paradata) from these activities were analysed to derive the cost savings from EQ.

33. Often the EQ mode is added to existing collection modes. This meant examining the cost of collection as it was before EQ and after EQ implemented. Therefore, there are a number of variables that need to be considered when estimating any efficiency: the collection modes available, the sample size, the EQ take-up rate (corresponding reduction in mail-out units and data capture), the number of EQ bounce-backs, the EQ response rate and the amount of NRFU and FEFU that would be reduced. Other considerations are: technical support from both a respondent and collection point of view. These elements were included in the efficiency estimation template and can be utilized as needed. The difference between the two estimates (before and after the introduction of EQ) represents the expected efficiencies.

C. EQ Experimental Design in the Unified Enterprise Survey

34. As the EQ mode of collection was relatively new at Statistics Canada, many questions were asked as to the best collection strategy that would lead to cost efficiencies while maintaining response rates and reducing respondent burden. Some of the questions were: Which follow-up strategy produces the best response rate? Which follow-up strategy is the most effective? How far can we get with e-mail reminders only? What is the importance of having the first contact via phone vs. e-mail? In order to establish standard collection follow-up strategy for electronic questionnaires, an experimental design was built using the UES 2011 EQ collection to compare different non-response follow-up (NRFU) methods (combining fax, telephone and e-mail reminders). Seven surveys in EQ were part of this study with almost 4500 units. Each treatment represents a different non-response follow-up approach using a combination of e-mail reminders, phone contacts and answering machine messages. The study was conducted in 2011. The table 6 below describes the treatments.

Table 6: UES Treatments

	T1 (Standard)	T2 (E-mails only)	T3 (phone attempt)	T4 (phone attempt + e-mails)
Reminders	E-mail only	E-mail only	1 st attempt – phone; 3 next – E-mail	1 st attempt – phone; 3 next – E-mail
Reminder schedule	1 st April 26 2 nd May 26 3 rd June 23 4 th July 25	1 st April 26 2 nd May 9 3 rd May 26 4 th June 7	1 st April 26-29 2 nd May 26 3 rd June 23 4 th July 25	1 st April 26-29 2 nd May 9 3 rd May 26 4 th June 7
Phone follow-up	Yes	No (firm appt only)	Yes	No (firm appt only)
Temporary follow-up flag	2	3	4	5
Experiment end date	August 1	June 20	August 1	June 20
Sample size allocated to the treatment	Appr. 25% of EQ sample	Appr. 25% of EQ sample	Appr. 25% of EQ sample	Appr. 25% of EQ sample

35. A comprehensive analysis was done on the experimental design, however this paper will provide some of the highlights. An analysis of the return rate as of June 20th was completed which marked the date for which all non-response units were switched to follow-up by telephone, regardless of their treatment. On June 20 (Table 7), with the exception of Treatment 2 (e-mails

only), all others treatments have similar return rates. 42% of the questionnaires were returned by sending only e-mail reminders. It is even more surprising to see that Treatment 4, where only one phone attempt was made between April 26 and April 29, attained almost the same results as Treatment 1 (standard treatment) and as Treatment 3, where telephone follow-ups were applied continuously since April 26.

Table 7: Questionnaires received by treatment (on June 20, 2011, end of experiment for T2 and T4)

	T1			T2			T3			T4		
	#	%w	%u	#	%w	%u	#	%w	%u	#	%w	%u
Total	731	49.6	47.8	651	42.2	41.8	766	47.8	49.6	749	46.4	48.3

The column “#” gives the number of questionnaires received and/or completed.

The column “%w” gives the weighted return rates.

The column “%u” gives the unweighted return rates.

36. After June 20th, a similar collection approach was done for all treatments including telephone follow-up along with e-reminders repeated at different intervals. At the end of collection on October 14 (Table 8), the four NRFU treatments all had very similar return rates.

Table 8: Questionnaires received treatment (on October 14, 2011)

	T1			T2			T3			T4		
	#	%w	%u	#	%w	%u	#	%w	%u	#	%w	%u
Total	1,125	84.0	81.8	1,102	81.9	81.4	1,148	80.3	82.4	1,114	83.0	81.0

37. The progression of total unweighted return rate over time by treatment is shown in Table 9. At the beginning of the non-response follow-up (April 26), the unweighted return rate of the four treatments are similar. On May 9, Treatments 2 and 4 received an e-mail reminder but Treatments 1 and 3 did not receive one. It is clear that after May 9, Treatment 2 and especially Treatment 4 had moved ahead in terms of return rate over the other two treatments. The trend continues until the end of May for Treatment 2 and until mid June for Treatment 4. It seems that having a reminder action every two weeks instead of every month is beneficial. On June 7, Treatment 2 received its fourth e-mail reminder. This fourth e-mail reminder did not seem to have a similar positive impact as the three previous ones. By the end of collection, the four NRFU treatments show very similar return rates.

Table 9: Progressive unweighted return rate (%) for each treatment (2011)

	T1	T2	T3	T4
	%u	%u	%u	%u
April 26	13%	12.5%	15%	12%
May 9	24%	22%	24%	22%
May 26	36%	41%	37%	45%
June 7	44%	43%	46%	49%
June 20	51%	46%	53%	52%
August 1	66%	54%	68%	58%
Sept 12	77%	76%	78%	75%

38. The number of attempts and the length of NRFU attempts amongst the four treatments were analysed. A measure of cost was derived by calculating the number of attempts necessary to obtain a response. The following table (Table 10) shows the average cost per respondent expressed in terms of number of call attempts by treatment (on October 14, 2011). The study shows that Tier 2 has the most cost effective collection strategy.

Table 10:

	T1		T2		T3		T4	
	w	u	w	u	w	u	w	u
Total	3.6	3.8	2.8	2.8	4.6	4.3	3.8	3.9

39. The following conclusions can be drawn from the study:
- (a) Electronic questionnaire collection treatments obtained a higher response rate, faster response time and shorter editing time than mail-out/mail-back collection method;
 - (b) NRFU treatments consisting of only e-mail reminders (or one phone attempt and then e-mail reminders) had enabled collection to obtain between 45% to 52% of questionnaires returned with very few phone calls;
 - (c) The four NRFU treatments show very similar response rates, but treatment T2 (email only) showed a relative cost lower than the other three treatments that show similar relative costs

D. Access Code Letters to Introduce EQ

40. In the spring of 2012, another test was conducted to assess the success level using an access code letter rather than mailing a questionnaire to respondents for whom pre-contact was not achieved. The first phase of the annual collection cycle for Unified Enterprise Survey (UES) consists of a telephone pre-contact for new businesses selected in the sample to confirm their contact information as well as the main business activity. It allowed identifying units out of scope of the survey due to changes in the industry activity or cessation of activity before starting collection of the survey proper. Pre-contact process also determined the initial mode of collection to be used. For EQ surveys, a telephone pre-contact was conducted to confirm existing information, to advise that collection would be done via electronic questionnaires, and to obtain e-mail addresses of the businesses.
41. The second phase of the collection consists of mailing out either paper questionnaires or e-mail invitations to the sampled units. In 2012, a new mail-out strategy was adopted for units which collection could not obtain the e-mail address during pre-contact. Instead of e-mailing an EQ invitation only to businesses that accepted to give their e-mail address during pre-contact, EQ invitations were sent to all businesses that have an e-mail address on the BR. Moreover, units unable to make contact during pre-contact period received an access code letter by mail that contained a hyperlink to an EQ. The only businesses assigned to paper questionnaires were those who requested a paper questionnaire specifically during pre-contact.
42. It should be noted that for those businesses that couldn't be reached during pre-contact and for which an e-mail address was not available on the BR, there is a higher risk of being non-respondents. Therefore, it is anticipated that lower response rates and higher effort will be spent for these businesses. Sample units who initially received an access code letter also received phone follow-up one month after the mail-out, like units in mail collection stream who were sent a paper questionnaire. Note that businesses that received an access code letter had the opportunity to request a paper questionnaire when they were contacted for follow-up.
43. Table 11 shows the status of collection in mid-July (collection ends on October 5th) for businesses that received an access code letter. A total of 765 units received an access code letter. Many are still in progress but of those who have responded, 79% of respondents completed the e-questionnaire and 21% responded by paper questionnaire. There is also more non-response for units that received an access code letter (7%) compared to email and mail collection (approx. 0%). As noted previously, businesses that received an access code letter are mostly new units in

the sample that collection staff could not reach during pre-contact and have no e-mail address already on record. It's not surprising to observe lower return rates and high out of scope and refusal rates for these cases.

Table 11: Collection status on July 16, 2012 for units that received an access code letter

	Status					
	Access code letter					
	# units (total)	Resp. (EQ)	Resp. (Mail)	Non-Resp.	Out of Scope	In Progress
Total	765	150 (20%)	39 (5%)	52 (7%)	74 (9%)	450 (60%)

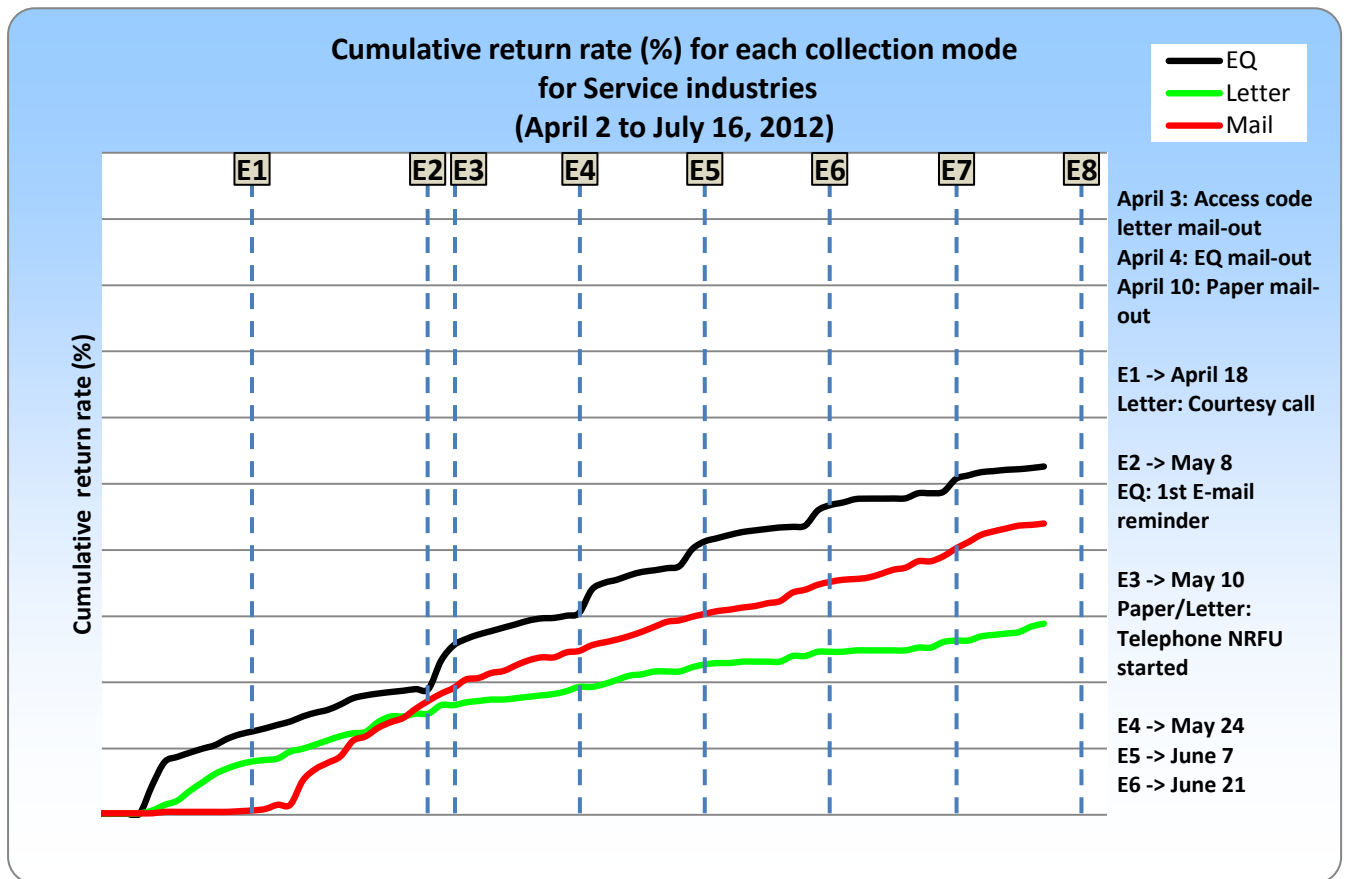
44. The return rate is presented for July 16 in Table 12. On July 16, return rate for EQ collection was somewhat higher than for mail collection and significantly higher than for access code letter.

Table 12: Questionnaires received by survey and collection mode on July 16, 2012

	EQ		Access code letter		Mail	
	#	%	#	%	#	%
Total	3,441	53.1	189	27.4	540	48.0

45. The progress over time of the cumulative percentage of completed questionnaires received for each collection mode is shown in Graph 1. Collection of units receiving an access code letter was not as successful as EQ (as of July 16th, the response rate using the access code letter was 29%; 44% for those who received a paper questionnaire by mail; and 52% for EQ). Nevertheless, when considering that contact was not made during the pre-contact, the results for the access code letter are promising. Almost 30% of respondents returned a completed e-questionnaire with minimal collection effort.

Graph 1:



46. The fact that the access code letter strategy doesn't seem to perform as well as the other two groups doesn't determine if it's an efficient strategy or not. The comparability of the three groups is limited, and the access code letter group was expected to yield lesser results. Looking at the final collection mode used by the few respondents of this group, it was observed that so far about 80% responded online using the EQ application. This seems to indicate that the strategy is working fairly well in encouraging the use of EQ, but a more serious analysis of the different costs involved would help determine if this strategy is actually cheaper than sending paper questionnaires right away instead of access code letters. An analysis comparing the cases from last year for which contact was not made during pre-contact (who received a paper questionnaire at the time) to those this year would be more relevant.

V. The future

47. The e-questionnaire project has been very successful in achieving its objectives in the first two years of its existence – a corporate e-questionnaire platform was built; surveys have been migrated to e-questionnaire as per the project deployment schedule; and targeted take up rates and efficiencies were achieved. At the same time, the project continues to evolve in response to new challenges and new business needs.
48. For business surveys, the next two major steps for the project are to begin e-questionnaire collection for mission-critical monthly surveys and to move the entire Unified Enterprise Survey program into the e-questionnaire world (60 surveys). To achieve these objectives, the project will need to add important new functionalities to the platform, such as ability to attach files to an e-questionnaire, self-coding tools for respondents (NAICS) or a solution for large enterprises to upload their data into e-questionnaires. On the questionnaire design front, the project will explore and test a possibility to further increase a number of data quality edit checks and will attempt to assess the level of tolerance of respondents to the edit checks. In regards to EQ survey implementation, there were many lessons learned that can be applied to the next surveys that will move to EQ. Work continues to automate processes, integrate systems, streamline the operations, and improve documentation and MIS. Efforts will continue to evaluate and monitor

costs and response rates. Research and continued enhancements are still required to optimise collection strategies.

49. For household surveys, the work has just begun to explore different multi-mode collection strategies that would fit with cross-sectional and longitudinal surveys that use multiple frames. Several pilots are scheduled to take place in 2012-13 to provide quantitative information for future directions. Based on the pilot results, we will work towards solidifying a multi-mode collection strategy with the use of e-questionnaire for different types of household surveys.
50. E-questionnaire design is another area that will see further improvements in the near future. With e-questionnaire proven to be a viable primary collection mode, design of questionnaires and collection tools has to be revisited to adapt to this new mode. Roles and responsibilities as well as processes and procedures need to be further reviewed and optimized.
51. Finally, Statistics Canada has several other important projects on the go. Some of these projects will have a significant impact on the e-questionnaire project and the e-questionnaire project will need to adapt to upcoming changes while continuing to integrate the e-questionnaire collection mode in our surveys as per the project deployment schedule and the efficiency targets.

VI. References

Burgess, J. (2007) Everything you wanted to know about Electronic Data Reporting (EDR) at Statistics Canada. Third International Conference on Establishment Surveys, 2007. Available at <http://www.amstat.org/meetings/ices/2007/index.cfm?fuseaction=confinfo>

Claveau, J., Leung, J., Turmelle, C. (2011) Design of Experiment for Non-response Follow-up of E-Questionnaire Collection

Claveau, J., Leung, J., Turmelle, C. (2012) Analysis of RY2011 New Electronic Questionnaire (EQ) Non-response Follow-up Strategy

Faid, M. (2009) Statistics Canada New e-Collection Services: Vision, Approach, Results. Information Technology Conference, 2009. Available at <http://www.statcan.gc.ca/cgi-bin/af-fdr.cgi?l=eng&loc=/conferences/it-ti2009/ppt/session19-faid-eng.ppt>

Karaganis, M. (2011), "Development and Implementation of E-questionnaire as a Primary Collection Mode at Statistics Canada", Internal document, Statistics Canada.

Karaganis, M., Fox, K., Claveau, J., Leung J. and Lin W. (2011), "Embedded Experiment for Non-response Follow-up Methods of Electronic Questionnaire Collection", Statistics Canada International Symposium Series, Proceedings, Statistics Canada, Cat. No. 11-522-X.

Karaganis, M. and Laflamme, G. (2012) "Use of e-questionnaires in household surveys", Technical report presented at Statistics Canada's Advisory Committee on Statistical Methods, April 2012.

Mayda, J. (2002) Experience with Implementation of EDR into Existing Survey Programs. *Statistical Journal of the United Nations Economic Commission for Europe*, 19(3), 131-140. Available at www.unece.org/stats/documents/2002.02.edr.htm.

Statistics Canada (2011), "E-Questionnaire Design Guidelines and Standards", Internal document, Ottawa, Canada.

Statistics Canada (2012), "Statistics Canada E-Questionnaire Collection Systems – Examples of screenshots", Internal document, Ottawa, Canada.

Appendix I: Collection strategies for multi-modal annual business surveys

	Date	Days after Mail out	Type of FU
Pre-contact		n/a	
Mail-out / E-invitation send date:		0	
Collection Start date:			
Bounce-backs follow up		2	1 week max to complete then switch to mail for those unable to get correct email
1st Follow up attempt date: Phone		30	only for those without e-mail address (mail units)
EQ		30	1st reminder email
Fax		X	no fax unless requested
Mail		X	no mail unless requested
2nd Follow up attempt date: Phone		60	only for those without e-mail address (mail units)
EQ		Every 2 nd week	2nd reminder email
Fax		X	no fax unless requested
Mail		X	no mail unless requested
3rd Follow up attempt date: Phone		90	only for those without e-mail address (mail units)
EQ		Every 2 nd week	2nd reminder email
Fax		X	no fax unless requested
Mail		X	no mail unless requested
4th Follow up attempt date: Phone		120	only for those without e-mail address (mail units)
EQ		Every 2 nd week	3rd reminder email
Fax		120	fax reminder + copy of questionnaire to those without email address
Mail		120	mail only to those without email address or fax number
5th Follow up attempt date: Phone		140	all outstanding
EQ		Every 2 nd week	5th reminder email
Fax		140	fax reminder + copy of questionnaire to those without email address
Mail		140	mail only to those without email address or fax number
6th Follow up attempt date: Phone		160	all outstanding units
EQ		105	6th and final reminder
Fax		X	no fax
Mail		X	no mail