Multivariate Selective Editing in the Integrated Business Statistics Program

Fraser Mills, Serge Godbout, Keven Bosa, Claude Turmelle
100 Tunney's Pasture Driveway, Ottawa ON, K1A0T6

Abstract

Statistics Canada is undergoing a redesign of its business surveys. One key component of the new framework is the adaptive selective editing methodology. Using historical and partially collected data, estimates and quality indicators are produced while collection is still underway. Item scores are calculated in order to gauge a unit's impact with regard to the quality indicator. These scores are then aggregated within each collection unit, creating a global unit score. Based on these, decisions regarding selective editing will be made, including producing priority lists for follow-up.

This talk will describe the adaptive selective editing methodology with quality indicators as focal points as well as the strategy proposed to integrate sampling, active collection and selective editing. Empirical results and potential savings in the new integrated business program will be discussed.

Key Words: Business Surveys, Adaptive Design, Quality Indicators, Selective Editing

1 Introduction

1.1 Context

The Corporate Business Architecture (CBA) was created to help Statistics Canada reduce costs while maintaining high quality services and improving timeliness and responsiveness of its statistical programs. The Integrated Business Statistics Program (IBSP) was proposed as one way of achieving CBA's objectives for business statistics. The IBSP will transform the current platform for producing annual business statistics; the Unified Enterprise Surveys (UES). It will also involve a larger array of surveys than the UES does.

One of the main goals of the IBSP is to achieve greater efficiency in processing its survey data, while producing estimates of similar, if not better, quality. To do this, a new adaptive design has been developed to manage data collection activities as well as data analysis (Godbout et al. (2011); Godbout (2011)). The Rolling Estimates (RE) model is a processing strategy that combines active collection management, editing, imputation, estimation and analysis. It allows estimates to be produced periodically as soon as an acceptable amount of survey and administrative data is available. Collection stops if all quality targets of a survey are met, increasing the timeliness of annual estimates and reducing the amount of resources dedicated to manual editing.

The original description of the RE model was described in Godbout et al. (2011). A more detailed treatment of the methodology surrounding the process was presented in Turmelle et al.

(2012). This paper seeks to provide the latest updates to the methodology as well as certain results from an empirical study which was undertaken.

1.2 The Current UES Survey Processing Model

Currently, the processing model of Statistics Canada's business surveys is typically linear (Figure 1). As the data come in through collection they are partially edited and some businesses are followed-up for non-response (based on weighted response rates) or failed edits (based on various edit rules). The complete set of data is then fully edited and imputed, and then analyzed by subject matter analysts. Finally, estimates are produced, analyzed and disseminated.

This is carried out through a long series of processes that require considerable manual intervention, even to run its automated steps. The time from processing to dissemination can be up to 6 months, and quality indicators regarding the estimates are only available during final analysis.

Figure 1: Current survey processing model

While the current process model produces good quality estimates, it is very lengthy and still heavily focussed on micro-data. In addition, follow-up is prioritized based on frame information, not on the estimates themselves or their quality.

1.3 The IBSP Survey Processing Model

In the IBSP, an optimal use of the resources available will be reached by limiting manual processing to the more influential units. To achieve this, the estimates and their quality must be taken into account during collection and processing rather than only near the end. The UES currently manages collection through the use of weighted response rates. However, for the IBSP, collection efforts will be actively managed based on key estimates and their quality indicators.

To do this, the IBSP will implement a circular approach (Figure 2) called the Rolling Estimates (RE) model. In this model, once enough data from administrative sources and collection have been received, a series of automated processes will be run, right through to producing estimates and their quality indicators. This information will then be analyzed using a top-down - macro estimates first, micro-data second - approach. The current plan is to produce these "rolling estimates" about once a month during a 4 to 5 month period.

The RE will produce key estimates and related quality indicators. Using these results, decisions will be made whether to stop active collection or not. When the quality indicators have reached pre-specified targets for a given geography-industry domain, active collection can stop in that domain and resources can be redirected towards other domains, as required.

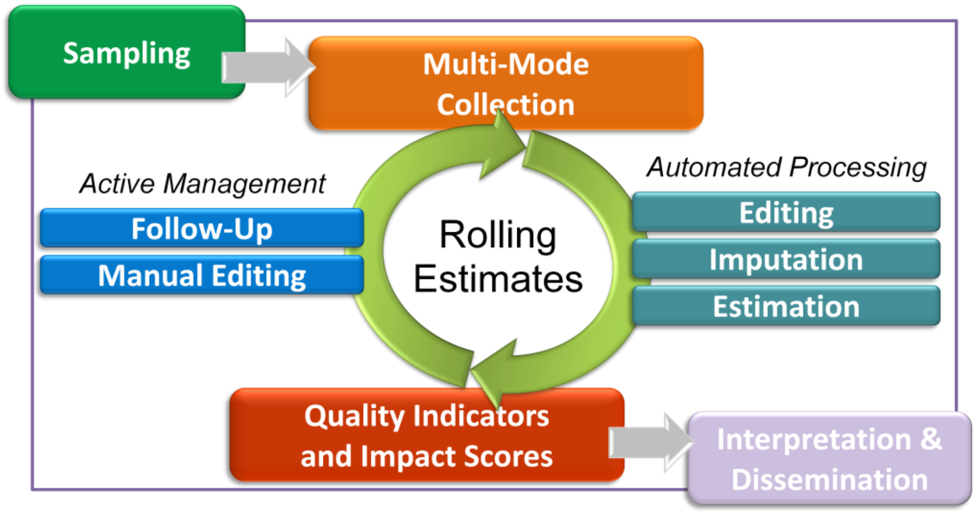

Figure 2: Rolling Estimates Model

The RE will also produce scores, for each unit, at each iteration. Active collection will be based on lists of non-responding units or units that failed collection edits, prioritized by their unit scores. Active analysis (also called Selective Editing) will mainly focus on respondents significantly influencing key estimates and their quality or non-respondents that are not eligible for collection follow-up (e.g. hard refusals).

2 Sampling and Estimation Methodology

We are interested in measuring a set of parameters $Y_j = \sum_{k \in U} y_{jk}$, for the combinations of variables of interest y_j ($j = 1, \dots, J$), for a population U of size N. A sample s of size n is drawn using a 2-phase design with stratified Bernoulli sampling at both phases. Unit or item non-response will be handled by imputation, defining the non-overlapping subset, $s_{jr} \cup s_{jm} = s$, as the set of k units having respectively reported $y_{jk}^{(R)} \equiv y_{jk}$ and imputed $y_{jk}^{(I)}$ values for the variable of interest y_{jk}. The set of respondents is identified as s_{jr}, and non-respondents, s_{jm}.

The estimator for the totals $\hat{Y}_j^{(IMP)}$ under imputation is given by:

$$\begin{aligned}\hat{Y}_j^{(IMP)} &= \sum_{k \in s_{jr}} w_k^{(E)} y_{jk}^{(R)} + \sum_{k \in s_{jm}} w_k^{(E)} y_{jk}^{(I)} \\ &= \sum_{k \in s} w_k^{(E)} y_{jk}^{*}\end{aligned} \tag{1}$$

The estimation weight $w_k^{(E)} = \pi_k^{-1} g_{ks}$ is the inverse of the sampling probability π_k resulting from the 2-phase design calibrated to the stratum counts and domain totals from administrative sources.

The IBSP methodological framework requires linear imputation methods i.e. the imputed value for a unit $k \in s_{jm}$, can be written in the linear form $y_{jk}^{(I)} = \varphi_{j0k} + \sum_{k' \in s_{jr}} \varphi_{jk'k} y_{jk'}^{(R)}$ where the quantities φ_{j0k} and $\varphi_{jk'k}$ do not depend on the *y*-values. Examples of linear imputation methods include auxiliary value imputation, linear regression and donor imputation. For more details, see Beaumont and Bissonnette (2011).

3 Active Management Framework

Some of the collection and analysis activities (referred to as treatments), like fax or email follow-ups, have a relatively low unit cost while other have significant marginal costs. The active management tries to balance the efforts from the high cost activities to produce data with quality corresponding to their uses. The high cost activities covered by this paper are telephone follow-up for non-response or failed-edit, and manual editing (Claveau et al. (2012)); they are all grouped into one single treatment *T*.

3.1 Key Estimates, Importance Factors and Quality Targets

The active management parameterization is done through the settings of 3 basic concepts: the list of key estimates, their importance factors and their quality targets. An estimate is identified from 3 attributes: a statistical measure, a variable (or many for multivariate statistical measures like ratios) and a domain. In the first years of IBSP, key estimates will consist of totals only.

All key estimates are assigned an importance factor ω_j, used to weigh their relative importance in the active management system, and a quality target QT_j to determine when the quality of an estimate is deemed sufficient for its use. The importance factor and the quality targets can be derived from an auxiliary data source or set manually.

3.2 Quality Indicator and Quality Distance

A quality measure Θ is a type of statistical measure θ used to assess the quality of an estimate or a set of estimates, with some common quality measures being the coverage rate, the response rates, the coefficient of variation (CV) and the relative root of the mean squared error (RRMSE). A quality indicator (QI) is the value, calculated or estimated, taken by a quality measure for a given estimate, i.e. $\hat{Q}I_j = \Theta(\hat{Y}_j | s, s_{jr})$.

For a decreasing quality indicator (i.e. being maximal before collection starts and going down to 0 when the quality has reached its target), the quality distance $\hat{Q}D_j$ for a given key estimate *j* is defined as $\hat{Q}D_j = \max\{(\hat{Q}I_j - QT_j) / \hat{Q}I_j, 0\}$. The objective is to ensure that the quality indicators for all key estimates meet or better their quality targets. This is similar to reducing all the quality distances to 0. To simplify the management of a set of quality indicators, it is convenient to combine them into a global statistic using a distance function (Hedlin, 2008). The distance function used in IBSP is a weighted quadratic mean derived from the multivariate objective function, also used at sampling (Turmelle et al., 2012). Therefore, the active management can be described by minimizing the global quality distance ($\hat{Q}D^G$), under constraints on costs, defined as:

$$\begin{aligned}\hat{Q}D^G &= \sqrt{\sum_{j=1,\ldots,J}(\omega_j \max(\hat{Q}I_j - QT_j, 0))^2 \Big/ \sum_{j=1,\ldots,J}\omega_j^2} \\ &= \sqrt{\sum_{j=1,\ldots,J}(\omega_j \hat{Q}D_j \hat{Q}I_j)^2 \Big/ \sum_{j=1,\ldots,J}\omega_j^2}\end{aligned} \tag{3}$$

The global quality distance is positive as long as there are key estimates for which the quality targets have not been met and decreases as their quality indicators improve; it will be 0 if and only if all the quality targets are met.

3.3 Unit Scores

The top-down solution used to go from the multivariate objective function from formula (3) to a collection unit prioritization is the measure of impact (MI) unit score. As described by Turmelle et al. (2012), the definition of the impact $\delta_{kT}(\hat{\theta}_j)$ of a treatment T on a unit k on a statistical measure $\hat{\theta}$ and a variable y_j conditionally to s and s_{jr}, is given by:

$$\delta_{kT}(\hat{\theta}_j) = \hat{\theta}_j - \tilde{\theta}_{j|kT} \tag{4}$$

Where $\tilde{\theta}_{j|kT}$ is the predicted effect on $\hat{\theta}_j$ of treatment T on unit k, assuming treatment will be successful. In the case a quality measure is a function of a vector of statistical measures, i.e. $\Theta = f(\boldsymbol{\theta})$ and $\hat{Q}I_j = f(\hat{\boldsymbol{\theta}}_j) = f(\hat{\boldsymbol{\theta}}(\hat{Y}_j | s, s_{jr})))$, the item score of the unit k, item j, under the treatment T is defined as $\hat{M}I_{jkT} = f(\boldsymbol{\delta}_{kT}(\hat{\boldsymbol{\theta}}_j))$. The unit score $\hat{M}I_{kT}^G$ of a unit k under the treatment T is given by:

$$\hat{M}I_{kT}^G = \sqrt{\sum_{j=1,\ldots,J}(\omega_j \hat{Q}D_j \hat{M}I_{jkT})^2 \Big/ \sum_{j=1,\ldots,J}\omega_j^2} \tag{5}$$

This unit score measures the impact a unit has on the global distance between the quality indicators and their targets. The score of a unit will be:

- Positive if it has a positive MI score for at least one key estimate for which the quality target has not been met yet;
- Zero if all its MI scores are zero or all its positive MI scores correspond to key estimates which have met their target.

3.4 Prioritization Lists

The unit scores are used to create a prioritization list for collection and analysis operations. Based on follow-up capacity, a threshold will be set to identify the priority units.

While there are no eligibility restrictions for analysis operations, not all units are eligible for collection follow-up because they have an appointment scheduled or they are excluded or excused. The priority units that are eligible for collection will be assigned to collection follow-up and the ineligible ones will be handled by analysis operations via selective editing.

4 Quality Measures

In this study, the active management will rely on the coefficient of variation making use of the total variance. In IBSP, a pseudo measurement bias will be added (Turmelle et al., 2012).

4.1 Total Variance Components

The total variance and its estimator are based on a methodology that requires the use of an imputation model (e.g., Beaumont and Bissonnette, 2011). We consider the following general imputation model:

$$\begin{aligned} &E_{\xi}\left(y_{jk} \middle| \mathbf{X}^{obs}\right) = \mu_{jk} \\ &\mathrm{var}_{\xi}\left(y_{jk} \middle| \mathbf{X}^{obs}\right) = \sigma_{jk}^{2} \\ &\mathrm{cov}_{\xi}\left(y_{jk}, y_{jk'} \middle| \mathbf{X}^{obs}\right) = 0, \end{aligned} \tag{2}$$

where $\mathbf{X}^{obs}$ is the observed matrix of auxiliary data, $k \neq k'$ and $k, k' \in U$.

As described by Beaumont et al. (2010), the total variance can be decomposed into the naïve sampling variance $\hat{V}_{Ord}$ term for which we consider the imputed values as reported, the correction $\hat{V}_{Dif}$ term, proposed by Särndal (1992) and simplified by Beaumont and Bocci (2009), to compensate for the effect of the imputation, the non-response variance $\hat{V}_{NR}$ term and the covariance $\hat{V}_{Mix}$ term. The estimation of the total variance $\hat{V}_{Tot}$ is the result of these four components:

$$
\begin{aligned}
&\hat{V}_{Tot}(\hat{Y}_j^{(IMP)}) = \hat{V}_{Ord}(\hat{Y}_j^{(IMP)}) + \hat{V}_{Dif}(\hat{Y}_j^{(IMP)}) + \hat{V}_{NR}(\hat{Y}_j^{(IMP)}) + \hat{V}_{Mix}(\hat{Y}_j^{(IMP)}) \\
&(1) \quad \hat{V}_{Ord}(\hat{Y}_j^{(IMP)}) = \sum_{k\in s}\sum_{k'\in s} \frac{\pi_{kk'} - \pi_k \pi_{k'}}{\pi_{kk'}} w_k^{(E)} e_{jk} w_{k'}^{(E)} e_{jk'} \\
&(2) \quad \hat{V}_{Dif}(\hat{Y}_j^{(IMP)}) = \sum_{k\in s_{jm}} (1-\pi_k) w_k^{(E)2} \hat{\sigma}_{jk}^2 \\
&(3) \quad \hat{V}_{NR}(\hat{Y}_j^{(IMP)}) = \sum_{k'\in s_{jr}} W_{jk'}^2 \hat{\sigma}_{jk'}^2 + \sum_{k\in s_{jm}} w_k^{(E)2} \hat{\sigma}_{jk}^2 \\
&(4) \quad \hat{V}_{Mix}(\hat{Y}_j^{(IMP)}) = 2\sum_{k'\in s_{jr}} W_{jk'} (w_{k'}^{(E)} - 1) \hat{\sigma}_{jk'}^2 - 2\sum_{k\in s_{jm}} w_k^{(E)} (w_k^{(E)} - 1) \hat{\sigma}_{jk}^2
\end{aligned}
\tag{6}
$$

where e_{jk} is a calibration residual obtained by treating the imputed values as true values. The quantity $W_{jk'} = \sum_{k\in s_{jm}} w_k^{(E)} \varphi_{jk'k}$ could be seen as the extra weight carried out by a unit k' with a reported value to compensate for the set of units imputed according to the model. More information about how those components of variance are derived can be found in Beaumont and Bissonnette (2011) and in Beaumont et al. (2010).

4.2 Item Score of Total Variance Components

The impact on the variance components are derived from formula (4) and (6). The effect of the treatment T on unit k consists of moving k from subset s_{jm} to s_{jr}. We assume that $W_{jk} = 0$ and that the treatment does not modify $y_{jk'}^*$, $\hat{\mu}_{jk'}$ and $\hat{\sigma}_{jk'}^2$, for $k' \in s$. Note that $\hat{\mu}_{jk'}$ and $\hat{\sigma}_{jk'}^2$ are consistent estimates of the fixed quantities $\mu_{jk'}$ and $\sigma_{jk'}^2$. The resulting total impact and component impacts are given by:

$$
\begin{aligned}
&\delta_{kT}(\hat{V}_{tot}(\hat{Y}_j^{(IMP)})) = \delta_{kT}(\hat{V}_{Dif}(\hat{Y}_j^{(IMP)})) + \delta_{kT}(\hat{V}_{NR}(\hat{Y}_j^{(IMP)})) + \delta_{kT}(\hat{V}_{Mix}(\hat{Y}_j^{(IMP)})) \\
&(1) \quad \delta_{kT}(\hat{V}_{Ord}(\hat{Y}_j^{(IMP)})) \approx 0 \\
&(2) \quad \delta_{kT}(\hat{V}_{Dif}(\hat{Y}_j^{(IMP)})) = (1-\pi_k) w_k^{(E)2} \hat{\sigma}_{jk}^2 \\
&(3) \quad \delta_{kT}(\hat{V}_{NR}(\hat{Y}_j^{(IMP)})) = w_k^{(E)2} \hat{\sigma}_{jk}^2 + \sum_{k'\in s_{jr}} \left(2W_{jk'} w_k^{(E)} \varphi_{jk'k} - (w_k^{(E)} \varphi_{jk'k})^2 \right) \hat{\sigma}_{jk'}^2 \\
&(4) \quad \delta_{kT}(\hat{V}_{Mix}(\hat{Y}_j^{(IMP)})) = -2 w_k^{(E)} (w_k^{(E)} - 1) \hat{\sigma}_{jk}^2 + 2\sum_{k'\in s_{jr}} w_k^{(E)} \varphi_{jk'k} (w_{k'}^{(E)} - 1) \hat{\sigma}_{jk'}^2
\end{aligned}
\tag{7}
$$

4.3 Quality Indicator and Item Score for IBSP

The quality indicator used to drive the active management in the IBSP and the resulting item score of unit k under treatment T based on the coefficient of variation are given by:

$$
\begin{aligned}
\hat{Q}I(\hat{Y}_j^{(IMP)}) &= \frac{\sqrt{\hat{V}_{Tot}(\hat{Y}_j^{(IMP)})}}{\hat{Y}_j^{(IMP)}} \\
\hat{M}I_{jkT}(\hat{Y}_j^{(IMP)}) &= \frac{\sqrt{\delta_{kT}(\hat{V}_{tot}(\hat{Y}_j^{(IMP)}))}}{\hat{Y}_j^{(IMP)}}
\end{aligned}
\tag{8}
$$

5 Rolling Estimates Empirical Study

In order to test the performance of the Rolling Estimates, a parallel run for the empirical study was designed to mimic the adaptive design method, using UES data from reference year 2011 for 46 different surveys. There were four iterations to the Rolling Estimates study (July, August, September, and October).

5.1 Parallel Run Processing

During the collection production of the UES, reference year 2011, a parallel run was carried out whereby all relevant micro- and macro- data related to editing, imputation, and estimation were stored at monthly intervals in order to simulate an automated processing system. This was accomplished using the UES tools with little additional processing elements. The resulting data was used in several testing and comparison studies related to the IBSP system requirements.

The UES design is not the same as the IBSP design in several ways. The following adjustments to our framework need to be taken into account when interpreting the results.

- Sampling is done in one phase in the UES, using a stratified simple random sample design, and calibration is not implemented in the UES' estimation strategy. The formula for $\hat{V}_{Ord}(\hat{Y}_j^{(IMP)})$ in (6) has been adjusted accordingly.
- This empirical study focussed on the non-response and not on data editing because the strategy in IBSP won't be comparable to that of the UES. The quality measure used in the study was the coefficient of variation. In IBSP, a pseudo measurement bias will be combined with the variance components to create a pseudo RRMSE.
- The collection and editing strategies and schedules will change: in UES, only six surveys have implemented an electronic questionnaire; in the IBSP, this will become the primary mode of collection and will affect the editing and collection procedures. The empirical study results for these six surveys were separated from the other ones.

5.2 The Empirical Study Procedure

The key variables and domains were a subset of those identified by the survey analysts in the UES and will not be the same in the IBSP. A subset of the UES key domains was used as the key estimates for the study. Specifically, the provinces and territories were used as the geography level, as well as national level. A partition based on the North American Industrial Classification System (NAICS) was used to define the industrial domains and certain surveys used one extra domain variable, like the for-profit status of the unit. The importance factors were defined in such a way that the provincial and territory totals for revenue had a very similar importance while the importance of the other variables and domains were set based on their relative provincial contribution. The quality targets were based on the final quality indicators estimated from the last iteration. The quadratic weighted average was taken and the individual targets were created based on this average and their importance factors.

At each iteration, the steps simulating non-response follow-up were as follows:

1) The quality indicators at a given iteration were estimated and compared with their targets. All the key estimates meeting their targets were identified; the impact of the affected item scores on the quality distance was eliminated.
2) For the key estimates that did not meet their target, the item scores were measured and combined into one score for each unit. Units only affecting key estimates that reached their target are zeroed out.
3) The top-N units with the largest unit scores were flagged for follow-up. Assuming resolution of these units, their impact was applied to the quality indicators in the proceeding iterations.

There were some important assumptions that affected the precision of the results. First, it was assumed that all units flagged in the prioritization lists were respondent by the next monthly iteration. In reality, there will be cases of flagged units remaining non-respondent or unresolved before the next iteration is run. Secondly, all units were assumed eligible for non-response follow-up. However, in practice eligibility flags would streamline collection and editing processes, so that, for instance, a unit may be sent directly to selective editing.

5.3 Summary of results

The results presented in Table 1 refer to percentages of attainable targets. Attainable targets are defined as those key estimates with the sampling portion of the CV below the target. If the sampling portion is too large, the non-response portion of the CV can never be reduced enough through collection to meet the target. There were approximately 10% of key estimates in the parallel run that were deemed unattainable. Given that, in the IBSP, the sampling strategy will be aligned with the active management strategy, at least in terms of domains' relative importance. This was not the case in the parallel run and empirical study, therefore a reduction in the occurrences of unattainable targets will likely be observed in the IBSP.

Ideally, 100% of attainable targets would be met by the end of collection. The results in Table 1 show between 97% and 99% of targets being met in the Rolling Estimates empirical study because a relatively large number of units were required in order to meet the targets of the last few key domains due to the fact that the top-N largest unit scores were followed-up. The mechanic of the Rolling Estimates is able to achieve 100% targets, but does so much more efficiently when the rolling estimates are run more frequently. This is especially true nearing the end of collection.

Table 1: Results by Survey Group

Survey Groups	Number of Key Estimates	Percentage of attainable quality targets met			Percentage of collection units followed up for non-response		
		UES production		RE Empirical Study	UES production		RE Empirical Study
		July	October	October	Before July	July-October	Overall reduction
TOTAL (46 surveys)	**8,600**	**76%**	**85%**	**98%**	**47%**	**53%**	**34%**
Non-EQ* (40 surveys)	7,600	76%	85%	98%	51%	49%	31%
EQ* (6 surveys)	1,000	71%	82%	99%	22%	78%	49%

From Table 1, we can observe that at the first iteration in July, with only 47% of the collection units being followed up for non response during UES production, 76% of the quality targets were already reached. Between July and October UES production, which flagged 53% of the collection units for non-response follow-up, brought only an additional 9% of key estimates below their targets.

The Rolling Estimates study created four successive prioritization lists, one per iteration, containing the collection units having the largest impact on global quality distance. The combination of the prioritized collection units from these four iterations and the collection units followed up before the first iteration contains 34% less units than the set of all collection units followed up for non-response in the UES production. If the units flagged in the four prioritization lists would have been all successfully resolved, 98% of the quality targets would have been reached. In practice, it's not all the prioritized units that will be successfully convert to a respondent. The unresolved cases could be analyzed by a subject matter who may contribute useful auxiliary data which would reduce the score of the unit.

The results are even stronger for the 6 surveys using electronic questionnaires, with a reduction of 49% of the number of collection units prioritized for non-response follow-up. This can be explained by the new follow-up procedures implemented for the surveys using electronic questionnaires: the early telephone follow-up activities are replaced by email reminders, without significant impact on the response progress (Claveau et al., 2012). Because there was a higher proportion of telephone follow-ups done after the first Rolling Estimates iteration, the potential for relative savings is more important.

6 Conclusions

The key features of this innovative active management strategy are the dynamic Rolling Estimates model driven by improved quality indicators.

The quality indicators, combining sampling and imputation variances, give an accurate picture of the current quality of all the key estimates. With the importance factors and the quality targets, the framework converts the multivariate objective into a univariate problem, allowing the IBSP to more easily, and efficiently, manage collection and analysis activities. Also, the unit scores integrate the top-down approach by allocating the quality indicators to the microdata.

The Rolling Estimates model provides regular, relevant and output-oriented pictures of the quality progress and, in a timely fashion, identifies the estimates that meet the targeted quality so the resources are dynamically redirected to focus on the remaining estimates.

This study has shown that, at a given collection follow-up capacity, the Rolling Estimates model with a unit prioritization based on quality indicators and unit scores can improve the quality management strategy while also reducing significantly the number of follow-ups required to meet quality targets. However, due to the limitations of the study, the achieved, theoretical reductions compared with the UES results cannot be blindly transposed to IBSP. The level of savings that can be expected depends on the collection strategy, the desired level of quality, and the choice of key estimates and their relative importance.

The empirical study demonstrated the feasibility and the power of the model, but also highlighted major requirements on the collection, analysis, processing, and methodology services and on their interactions. The strategy, parameterized through key estimates, importance factors and quality targets, has to be carefully set up. Collection progress monitoring needs to be reviewed so that the joint efforts between collection staff and analysts may maximize non-response and edit resolutions. The Rolling Estimates model also requires enhancement in terms of system performance, robustness, flexibility, and method standardization and coherence.

The Rolling Estimates model using the quality indicator and the active management methodology will be implemented for the first year of IBSP. The plan is to start with reasonable expectations for estimate quality and follow-up savings then use the experience from the first few years to assess the efficiency of the framework under the IBSP model.

References

Beaumont, J-F. and Bissonnette, J. (2011). Variance estimation under composite imputation: The methodology behind SEVANI. Survey Methodology, Vol. 37, 171-180.

Beaumont, J-F., Bissonnette, J. and Bocci, C.(2010). SEVANI, version 2.3, Methodology Guide. Internal report, Methodology Branch, Statistics Canada.

Beaumont, J.-F., and Bocci, C. (2009). Variance Estimation when Donor Imputation is Used to Fill in Missing Values. *Canadian Journal of Statistics*, 37, 400-416.

Claveau, J., Leung, J. and Turmelle, C. (2012). Embedded Experiment for Non-Response Follow-Up Methods of Electronic Questionnaire Collection. *Proceedings of the ICES IV Conference.*

Godbout, S., (2011). Normalisation du traitement des données après la collecte dans les enquêtes-entreprises à Statistique Canada. Symposium international de 2011 sur la méthodologie. Associated document.

Godbout, S., Beaucage, Y. and Turmelle, C. (2011). Quality and Efficiency using a Top-Down Approach in the Canadian's Integrated Business Statistics Program. Conference of European Statisticians, Work Session on Statistical Data Editing. Ljubljana, Slovenia (9-11 May 2011).

Hedlin, D. (2008). Local and Global Score Functions in Selective Editing. United Nations Statisticl Commision and Economic Commission for Europe, Conference of European Statisticians, Work Session on Statistical Data Editing, Vienna, Austria 2008. Associated document.

Särndal, C.E., Swensson, B. and Wretman, J. (1992). Model Assisted Survey Sampling. New York, Springer-Verlag, 694 p.

Turmelle, C., Godbout, S. And Bosa, K. (2012). Methodological Challenges in the Development of Statistics Canada's New Integrated Business Statistics Program. *Proceedings of the ICES IV Conference.*