THE SAMPLE DESIGN OF THE QUARTERLY SURVEY OF FINANCIAL STATEMENTS: REVIEW AND IMPROVEMENTS

Nora Bohossian, Pierre Daoust, Sylvie Gauthier[1]

ABSTRACT

Statistics Canada's Quarterly Survey of Financial Statements collects balance sheet and income statement data for incorporated businesses in Canada. Its main objectives are to measure financial performance and activity by various industry aggregations in the Canadian economy and to provide business sector profit and financial flow inputs into Canada's System of National Accounts. The survey consists of a Take-All stratum of the largest enterprises and a sample of medium-size enterprises. Recently, various revisions were made to the sample design with the aim of achieving greater sample stability. In this paper, we present an overview of the design and discuss future improvements.

KEY WORDS: Financial Survey Methodology; Stratified Sample Design.

RÉSUMÉ

Le Relevé trimestriel des états financiers de Statistique Canada recueille de l'information sur les états des résultats et les bilans des entreprises canadiennes constituées en sociétés. Ses principaux objectifs sont de mesurer la situation financière et le rendement financier pour différents agrégats de l'économie canadienne et de fournir les renseignements sur les bénéfices du secteur des entreprises et sur les flux financiers pour le Système canadien de comptabilité nationale. Cette enquête est composée d'une strate à tirage complet des grandes entreprises et d'un échantillon d'entreprises moyennes. Récemment, divers aspects de son plan de sondage ont été revus afin d'augmenter la stabilité de son échantillon. Cet article présente un aperçu de ce plan et se penche sur les améliorations qu'il serait possible d'y apporter.

MOTS CLÉS : Méthodologie des enquêtes financières; Plan de sondage stratifié.

1. INTRODUCTION

The Quarterly Survey of Financial Statements (commonly referred to as the Quarterly Financial Survey or QFS) is a key survey at Statistics Canada. The data collected by this survey of incorporated businesses comprise financial variables obtained from corporate income statements (elements of revenues, expenses and profits) and balance sheets (elements of assets, liabilities and equity). Its outputs are used to track the performance of the business sector by industry aggregates, calculate business sector GDP and financial flows (inter-sectoral savings and investments and lending and borrowing), provide inputs into the national balance sheet and flag foreign financing activity.

The business sector is one of the four main sectors in the Canadian System of National Accounts (CSNA). The other sectors: household, government and the foreign sector are covered by other survey programs at Statistics Canada and are excluded from QFS. The data collected by QFS and the data collected from different sources covering the CSNA sectors are then combined to produce periodic accounts of the economic activity in Canada.

The survey frame is extracted from Statistics Canada's Business Register (BR) which is a list of businesses engaged in the production of goods and services in Canada. The BR represents businesses through their legal and operating structures. The survey target for the QFS is the enterprise, the highest legal unit in a corporate structure that can provide consolidated

[1] Nora Bohossian, 100 Tunney's Pasture Driveway, Ottawa, Canada, K1A 0T6, Nora.Bohossian@statcan.gc.ca; Pierre Daoust, 100 Tunney's Pasture Driveway, Ottawa, Canada, K1A 0T6, Pierre.Daoust@statcan.gc.ca; Sylvie Gauthier, 100 Tunney's Pasture Driveway, Ottawa, Canada, K1A 0T6, Sylvie.Gauthier@statcan.gc.ca

financial statements. There are roughly 2.8 million active enterprises on the BR but only about 1.2 million are in scope for QFS.

The units on the QFS frame are grouped by their economic activity using 80 industrial aggregations based on the North American Industry Classification System (Statistics Canada, 2007). Within each industry group, the units are further stratified by size using two size variables, revenue and assets. The stratum boundaries typically differ for each size variable and by industry group.

Within a typical industry group, large businesses form the Take-All (TA) stratum, medium businesses form the Take-Some (TS) stratum and small businesses form the Take-None (TN) stratum. However, there are exceptions. There are five census industries, that is, they consist of a TA stratum only. There are an additional nine industries which do not have a TS stratum: each business in these industry groups either belongs to the TA stratum or to the TN stratum.

Approximately 98% of all enterprises on the frame are small businesses (TN) representing 30% of the total revenue and 15% of the total assets. Although no units are surveyed from the TN stratum, a model is used to derive estimates for these units which are later combined with the estimates for the sampled units to obtain the final estimates. Throughout this paper, *sampled population* refers to the sampled portion of the survey population (that is, units in the TA and the TS strata) and *non-sampled population* refers to the non-sampled portion of the survey population (that is, units in the TN stratum).

Studies conducted in recent years on the distribution of revenue and assets of the TS units highlighted that in some industries the distribution is relatively skewed and spreads over a wide range of values. It was apparent that a better stratification approach would be to group the units that were considered large by their revenue or their assets into one TS stratum (TS1) and the remaining TS units into a second TS stratum (TS2). Such an approach was likely to create more homogeneous TS strata and reduce the sampling variability of the survey estimates.

The properties of the proposed modifications to the TS stratification were carefully evaluated. A pilot study investigated the new approach on a parallel run of the actual 2008 QFS sampling process. (Note that the original one-TS methodology was preserved for industries with a small number of TS-sized enterprises.) The pilot study suggested that at least 1% to 2% gains in the coefficient of variation (CV) could be expected for most of these industries while still retaining the same or similar sample size. (Most of the original CVs ranged between 3% and 9%.) Based on these promising results, the new stratification methodology was adopted and implemented in the first quarter of 2009 for the majority of the industry groups with a TS stratum. The observed improvements in the CVs were similar to those observed in the pilot study.

The main objective of this paper is to describe the QFS sampling and estimation strategy after the adoption of the new TS methodology. Section 2 gives a detailed overview of the QFS sample design. Alternative methods being considered for estimation of the survey's key variables are presented in Section 3. Future considerations and conclusions follow in Section 4.

2. SAMPLE DESIGN

Since there may be at most four size strata in each industry group, industries can be classified into one of the following four categories: (1) Industries that consist of one TA stratum only (a census); (2) Industries that consist of one TA and one TN strata; (3) Industries that consist of one TA, one TS and one TN strata; and (4) Industries that consist of one TA, two TS and one TN strata. In this paper, we will be discussing the methodology for the fourth category of industries, which can be easily simplified to apply to the other three categories.

The QFS sample design is different depending on whether it is a Refresh quarter, when the sampling properties can be altered for all enterprises in the sampled population, or an Update quarter, when most units retain their sampling properties from the previous quarter and the sampling of only a subset of the units (for example, new units) is considered. The following two sections describe the sample design for Refresh and Update quarters, respectively.

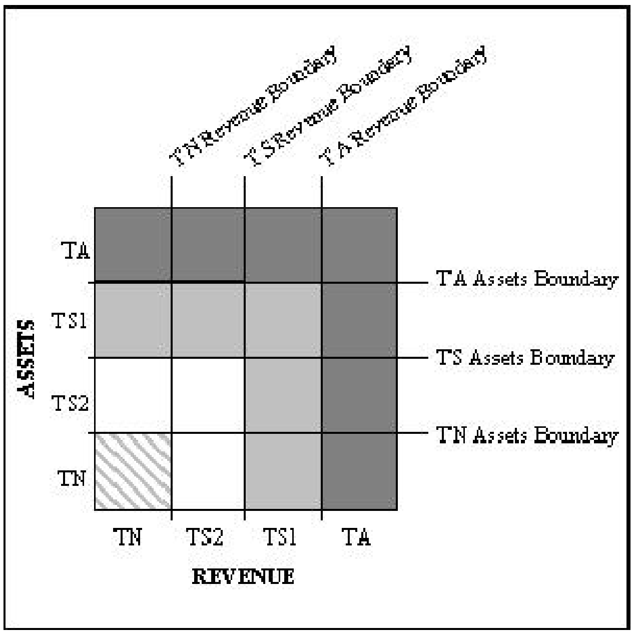
2.1 Refresh Quarter

The first quarter (Q1) of every calendar year is typically the Refresh quarter when the sample is adjusted for all industries. The TA and the TS boundaries for revenue and assets in each industry group are derived using a stratification algorithm based on the Lavallée-Hidiroglou method (1988). An adaptation of the method is first used to derive iteratively the TA boundaries and, after they have been determined, the Lavallée-Hidiroglou method is used to derive the TS boundaries separately for revenue and for assets which are then combined to produce the two TS strata.

The stratification algorithm determines the boundaries that yield the minimum required sample size while satisfying pre-specified constraints on the CV of total revenue and total assets estimates. The final sample size in each TS stratum must also satisfy a pre-specified minimum sampling fraction and a minimum sample size of four enterprises. Lastly, the TN boundaries are reviewed every Refresh quarter and are occasionally modified as a means of controlling the size of the sampled population which consists of roughly 21,000 enterprises.

Once the boundaries are finalized, the survey population is stratified as depicted in Figure 1 below. Note that units with more complex structures are stratified using pre-determined boundaries rather than the boundaries described above.

Figure 1 – Strata Definitions



In Figure 1 the patterned region represents TN units, the white region represents TS2 units, the light grey region represents TS1 units and, lastly, the dark grey region represents TA units. A unit is in a specific stratum if at least one of its two size variables is greater than or equal to the respective boundary for that stratum. For example, a unit is a TA if either its revenue or its assets is greater than or equal to the respective TA boundary. In the case of TS1/TS2 units, the TS boundary splits the TS stratum into TS1 (greater than or equal to the TS boundary) and TS2 (below the TS boundary).

In a Refresh quarter, the sample overlap with the previous quarter's sample is maximized using the Kish and Scott method (1971). All units of TA size are in the sample with certainty. Within each of the two TS strata a sample is selected. It is considered that it is being selected using Simple Random Sampling (SRS).[2] The remaining units in the QFS population comprise of the non-sampled portion of the survey population, the TN units. The sample size is roughly 5,500 units.

The subsequent three quarters (Q2, Q3 and Q4) after the Refresh quarter are Update quarters. The next sub-section describes the sample selection procedures in these quarters.

---

[2] It should be noted that this is not exactly the case because of a small number of new units that enter the sample in an Update quarter of the previous year and the Kish and Scott procedure being used at the Refresh quarter of the current year, which leads to a sample stratified on the intersection of the previous and current quarter strata.

2.2 Update Quarter

In an Update quarter, the survey frame is updated by extracting the most recent information available from the BR. As a result, some enterprises leave and new enterprises enter the survey population and, potentially, the sampled population. This means that only a subset of the units that were in the sampled population of the previous quarter is left, referred to as the *continuer* units. In order to account for this movement of units the sample needs to be updated.

Units that leave the sampled population of a particular industry group consist of three kinds: (1) Units that have moved from the sampled population to the non-sampled population; (2) Units that are no longer alive and active (and, hence, no longer on our frame); and (3) Units that changed industry group and are no longer classified under the same industry as in the previous quarter. The units in (1) and (2) are referred to as *dead* units with respect to the sampled population, whereas units in category (3) are continuer units since they continue to be part of the overall sampled population but just under a different industry group.

Moreover, units that enter the sampled population of a particular industry group also consist of three kinds: (1) Units that have moved from the non-sampled population to the sampled population; (2) Units that appear on the frame for the first time; and (3) Units that changed industry group and are new to the industry group in the current quarter. The units in (1) and (2) are referred to as *birth* units with respect to the sampled population, whereas units in category (3) are continuer units since they continue to be part of the overall sampled population but just under a different industry group.

All continuer units retain their sampling status from the previous quarter including the units that changed industries. The only exception arises for units which have moved to the TA portion of the survey population. Those units are in the current quarter's sample with certainty regardless of their sampling status in the previous quarter. To compensate for the units that have entered the population, the sample is updated by selecting units from the birth units in each of the two TS strata using Bernoulli sampling. This process is repeated every Update quarter until the next Refresh quarter.

Survey estimates are produced quarterly. The following section presents various estimation methods used to derive the estimates in an Update quarter. Estimation procedures for a Refresh quarter are straightforward and are not discussed in this paper.

3. ESTIMATION

The final estimates comprise two components, the estimates obtained from the sampled portion of the population (TA and TS units) and the estimates obtained from a model for the TN units (non-sampled portion). The latter is outside the scope of this paper and the current section discusses only the estimates derived from the sample survey.

We are interested in estimating the sampled population total, Y, for example total revenue or total assets, in a specific industry i for the current Update quarter. We begin by defining our sampling and estimation parameters.

The parameter j indicates the industry in which the unit was sampled, where

$$j = \begin{cases} \text{industry at current quarter, if birth unit} \\ \text{industry at previous quarter, otherwise} \end{cases}.$$

We also have the parameter h, which indicates the size stratum in which the unit was sampled within a given industry j. Essentially, h differentiates between TA, TS1 and TS2 units and, within the TS-sized units, h differentiates between continuer units that have been sampled using SRS and birth units that have been sampled using Bernoulli. Note that non-TA units which changed industry retain their sampling status from the previous quarter and are treated as continuer units even though they are new to the population of industry i. Thus h is defined as

$$h=\begin{cases}1, & if \quad TA\\ 2, & if \quad TS1DC\\ 3, & if \quad TS1B\\ 4, & if \quad TS2DC\\ 5, & if \quad TS2B\end{cases},$$

where, in the expression above: *TA* stands for units which belonged to the TA stratum in a Refresh quarter or in any one of the Update quarters; *TS1DC* stands for units which belonged to the TS1 stratum (larger TS-sized enterprises) in the Refresh quarter and are dead or continuer of TS size (TS1 or TS2) in the current Update quarter; *TS1B* stands for units which were birth units in the TS1 stratum in one of the Update quarters; *TS2DC* stands for units which belonged to the TS2 stratum (smaller TS-sized enterprises) in the Refresh quarter and are dead or continuer of TS size (TS1 and TS2) in the current Update quarter; *TS2B* stands for units which were birth units in the TS2 stratum in one of the Update quarters.

With these parameters, our sampled population in industry j and stratum h is denoted by S_{jh}. Our domain, d_i, consists of all units that are alive and active in industry i in the current quarter. Our population total is then given by

$$Y(d_i)=\sum_{j=1}^{80}\sum_{h=1}^{5}\sum_{k\in S_{jh}} y_{jhk}t_{jhk}(d_i)=\sum_{j=1}^{80}\sum_{k\in S_{j1}} y_{j1k}t_{j1k}(d_i)+\sum_{j=1}^{80}\sum_{h=2}^{5}\sum_{k\in S_{jh}} y_{jhk}t_{jhk}(d_i)=Y_{TA}(d_i)+Y_{TS}(d_i), \tag{3.1}$$

where $t_{jhk}(d_i)$ is the indicator variable that a unit is in the domain d_i and TS includes TS1 and TS2 units.

Since all units of TA size in our survey population are in our sample, we only need to find an estimator for $Y_{TS}(d_i)$. Let s_{jh} be the sampled enterprises in industry j and stratum h. Then, assuming the first quarter of the calendar year, Q1, is the Refresh quarter, we define the design weights, a_{jhk}, as follows

$$a_{jhk}=\begin{cases}1 & if \quad h=1\\ N_{j,Q1TS1}/n_{j,Q1TS1} & if \quad h=2\\ 1/f_{j,TS1} & if \quad h=3\\ N_{j,Q1TS2}/n_{j,Q1TS2} & if \quad h=4\\ 1/f_{j,TS2} & if \quad h=5\end{cases},$$

where for the j^{th} industry group: $N_{j,Q1TS1}$ and $N_{j,Q1TS2}$ are the population sizes of the TS1 and TS2 strata at the last Refresh quarter, respectively; $n_{j,Q1TS1}$ and $n_{j,Q1TS2}$ are the sample sizes of the TS1 and TS2 strata at the last Refresh quarter, respectively; $f_{j,TS1}$ and $f_{j,TS2}$ are the sampling fractions used at the current Update quarter for Bernoulli sampling of TS1-sized and TS2-sized birth units, respectively. Again, note that non-TA units, which changed their industry group in the current quarter, retain their weights from the previous quarter since their sampling status does not change.

Using these weights, the π estimator of the population total (Equation 3.1) is then given by

$$\hat{Y}_\pi(d_i)=\sum_{j=1}^{80}\sum_{h=1}^{5}\sum_{k\in s_{jh}} a_{jhk}y_{jhk}t_{jhk}(d_i)=Y_{TA}(d_i)+\sum_{j=1}^{80}\sum_{h=2}^{5}\sum_{k\in s_{jh}} a_{jhk}y_{jhk}t_{jhk}(d_i)=Y_{TA}(d_i)+\hat{Y}_{\pi,TS}(d_i). \tag{3.2}$$

However, this estimator could be inefficient due to variability of the sample size of the TS-sized continuer and birth units. To reduce the impact of this variability on the estimates, it is necessary to post-stratify on the current quarter strata counts.

The two approaches that are being considered, the combined ratio estimator and the separate ratio estimator, are described below.

3.1 Combined Ratio Estimator

The combined ratio estimator carries out the adjustment on the π estimator (Equation 3.2) by combining the two TS strata. In other words, we have

$$\hat{Y}_C(d_i) = Y_{TA}(d_i) + \hat{Y}_{C,TS}(d_i), \tag{3.3}$$

where

$$\hat{Y}_{C,TS}(d_i) = \frac{N_{TS}(d_i)}{\hat{N}_{\pi,TS}(d_i)} \hat{Y}_{\pi,TS}(d_i)$$

with $N_{TS}(d_i) = \sum_{j=1}^{80} \sum_{h=2}^{5} \sum_{k \in S_{jh}} t_{jhk}(d_i)$ and $\hat{N}_{\pi,TS}(d_i) = \sum_{j=1}^{80} \sum_{h=2}^{5} \sum_{k \in s_{jh}} a_{jhk} t_{jhk}(d_i)$.

A Taylor series expansion is used to approximate the variance of the combined estimator giving $V\left(\hat{Y}_C(d_i)\right) = V\left(\hat{Y}_{C,TS}(d_i)\right) \approx V\left(\hat{E}_{\pi,C,TS}(d_i)\right)$, where $\hat{E}_{\pi,C,TS}(d_i) = \sum_{j=1}^{80} \sum_{h=2}^{5} \sum_{k \in s_{jh}} a_{jhk} e_{C,jhk}(d_i) t_{jhk}(d_i)$ is the π estimator of the residuals

$$e_{C,jhk}(d_i) = y_{jhk} - \frac{Y_{TS}(d_i)}{N_{TS}(d_i)}. \tag{3.4}$$

The variance approximation above can be expanded as

$$V\left(\hat{Y}_C(d_i)\right) \approx V\left(\hat{E}_{\pi,C,TS}(d_i)\right) = V\left(\hat{E}_{\pi,C,TS1DC}(d_i)\right) + V\left(\hat{E}_{\pi,C,TS1B}(d_i)\right) + V\left(\hat{E}_{\pi,C,TS2DC}(d_i)\right) + V\left(\hat{E}_{\pi,C,TS2B}(d_i)\right),$$

consisting of the following four components:

(1) The variance of the dead and continuer units of TS1 size (which are a subset of the Q1 TS1 sample) given by

$$V\left(\hat{E}_{\pi,C,TS1DC}(d_i)\right) = \sum_{j=1}^{80} \frac{N_{j,Q1TS1}^2}{n_{j,Q1TS1}} \left(1 - {n_{j,Q1TS1}}/{N_{j,Q1TS1}}\right) S_{j,Q1TS1}^2\left(z_{C,jm}(d_i)\right) \tag{3.5}$$

with $z_{C,jm}(d_i) = e_{c,jm}(d_i) t_{jm}(d_i)$, where $e_{C,jm}(d_i) = e_{C,j2k}(d_i)$ for $k \in S_{j2}$ and $e_{C,jm}(d_i) = 0$ for $k \notin S_{j2}$, and $t_{jm}(d_i)$ is the indicator variable that the unit is in domain d_i. Moreover,

$$S_{j,Q1TS1}^2\left(z_{C,jm}(d_i)\right) = \frac{1}{(N_{j,Q1TS1} - 1)} \sum_{m \in S_{j,Q1TS1}} \left(z_{C,jm}(d_i) - \bar{z}_{C,j}(d_i)\right)^2$$

with $\bar{z}_{C,j}(d_i) = \sum_{m \in S_{j,Q1TS1}} {z_{C,jm}(d_i)}/{N_{j,Q1TS1}}$, where $S_{j,Q1TS1}$ refers to the population of the TS1 stratum in industry j at the Refresh quarter (including the units that move to the TA portion in the Update quarter).

(2) The variance of the birth units of TS1 size given by

$$V\left(\hat{E}_{\pi,C,TS1B}(d_i)\right) = \sum_{j=1}^{80} \sum_{k \in S_{j3}} (a_{j3k} - 1)\, e_{C,j3k}^2\, t_{j3k}(d_i). \tag{3.6}$$

(3) The variance of the dead and continuer units of TS2 size (which are a subset of the Q1 TS2 sample) given by Equation 3.5 with TS1 replaced by TS2 and $h = 4$.

(4) The variance of the birth units of TS2 size given by Equation 3.6 with TS1 replaced by TS2 and $h = 5$.

The formula used for components (1) and (3) could be simplified by conditioning on the number of sampled enterprises in the $TS1DC$ and $TS2DC$ domains that are alive and active in the current quarter for which the estimates are being derived. Under this conditional framework, the sample design of the corresponding sub-domains would be SRS.

The method described so far post-stratifies on the combined two TS strata. An alternative approach, described in the sub-section that follows, is to post-stratify on each of the two TS stratum separately.

3.2 Separate Ratio Estimator

We can define d_{TS1} as the domain of TS1-sized enterprises that are alive and active at the current quarter and, similarly, d_{TS2} as the domain of TS2-sized enterprises that are alive and active at the current quarter. (The domains differ from the defined strata because of stratum jumpers.). Then, we can break down the TS component of the π estimator (Equation 3.2) as follows

$$\hat{Y}_{\pi,TS}(d_i) = \hat{Y}_{\pi,TS1}(d_i) + \hat{Y}_{\pi,TS2}(d_i)$$

where

$$\hat{Y}_{\pi,TS1}(d_i) = \sum_{j=1}^{80} \sum_{h=2}^{5} \sum_{k \in s_{jh}} a_{jhk}\, y_{jhk}\, t_{jhk}(d_i)\, t_{jhk}(d_{TS1})$$

and

$$\hat{Y}_{\pi,TS2}(d_i) = \sum_{j=1}^{80} \sum_{h=2}^{5} \sum_{k \in s_{jh}} a_{jhk}\, y_{jhk}\, t_{jhk}(d_i)\, t_{jhk}(d_{TS2})$$

with $t_{jhk}(d_{TS1})$ and $t_{jhk}(d_{TS2})$ being the indicator variables that a unit is in domains d_{TS1} and d_{TS2}, respectively.

Then, we can use a separate ratio estimator which carries out the adjustment on the π estimator on each of the two TS components separately. In other words, we have

$$\hat{Y}_S(d_i) = Y_{TA}(d_i) + \hat{Y}_{S,TS1}(d_i) + \hat{Y}_{S,TS2}(d_i), \tag{3.7}$$

where

$$\hat{Y}_{S,TS1}(d_i) = \frac{N_{TS1}(d_i)}{\hat{N}_{\pi,TS1}(d_i)} \hat{Y}_{\pi,TS1}(d_i)$$

with $N_{TS1}(d_i) = \sum_{j=1}^{80} \sum_{h=2}^{5} \sum_{k \in S_{jh}} t_{jhk}(d_i)\, t_{jhk}(d_{TS1})$ and $\hat{N}_{\pi,TS1}(d_i) = \sum_{j=1}^{80} \sum_{h=2}^{5} \sum_{k \in s_{jh}} a_{jhk}\, t_{jhk}(d_i)\, t_{jhk}(d_{TS1})$, and

$$\hat{Y}_{S,TS2}(d_i) = \frac{N_{TS2}(d_i)}{\hat{N}_{\pi,TS2}(d_i)} \hat{Y}_{\pi,TS2}(d_i)$$

with $N_{TS2}(d_i) = \sum_{j=1}^{80} \sum_{h=2}^{5} \sum_{k \in S_{jh}} t_{jhk}(d_i)\, t_{jhk}(d_{TS2})$ and $\hat{N}_{\pi,TS2}(d_i) = \sum_{j=1}^{80} \sum_{h=2}^{5} \sum_{k \in s_{jh}} a_{jhk}\, t_{jhk}(d_i)\, t_{jhk}(d_{TS2})$.

The variance of the separate ratio estimator can also be approximated using a Taylor series expansion. In this case, the residuals are given by

$$e_{S,jhk}(d_i) = \begin{cases} y_{jhk} - \dfrac{Y_{TS1}(d_i)}{N_{TS1}(d_i)} & if \quad k \in d_{TS1}, \quad 0 \quad otherwise \\ y_{jhk} - \dfrac{Y_{TS2}(d_i)}{N_{TS2}(d_i)} & if \quad k \in d_{TS2}, \quad 0 \quad otherwise \end{cases}. \tag{3.8}$$

Then, the variance of the separate estimator, $V\left(\hat{Y}_S(d_i)\right)$, can be derived as the one for the combined estimator described in Section 3.1 by replacing $e_{C,jhk}(d_i)$ with $e_{S,jhk}(d_i)$ (that is, replacing Equation 3.4 with Equation 3.8).

As mentioned in the introduction, the new methodology was evaluated in parallel to the actual 2008 QFS sampling process. Both estimators were implemented and the CVs were compared to the CVs obtained with the original methodology. An improvement of 1% to 2% in the CVs was noted for the majority of the industries where the two-TS stratification was applied. The observed CV gains were similar for the two estimators with some indication that the separate estimator was slightly better than the combined estimator.

4. CONCLUSION

Results from the second quarter of 2009 showed that it is preferable to use the separate ratio estimator (Equation 3.7) over the combined (Equation 3.3). This is because in some industries movement of units is concentrated in one of the TS strata (usually in the smaller, TS2, strata) and adjusting on the two strata combined may introduce a conditional bias. However, the option of using the separate estimator for some industries and the combined estimator for the rest has not been eliminated.

As mentioned in the introduction, the QFS frame is extracted from the BR. The BR was recently redesigned allowing for timelier update of size measures data (such as revenue and assets). This has resulted in an increased number of enterprises changing strata/post-strata from quarter to quarter and more movement to, from and within the survey population. Particularly with the introduction of the second TS stratum, finding ways to reduce this movement has become quite important. Research studies have begun looking into the possibility of introducing boundary resistance rules that will stabilize the population over time by making the strata and post-strata definitions less sensitive to small but boundary crossing changes in the size variables.

Apart from the study on the boundary resistance rules, the possibility of implementing a partial Refresh or, in other words, redrawing the sample every Refresh quarter only for some industries as opposed to all is being investigated. One scenario could be to have a Refresh quarter every x number of years (currently, $x = 1$ so for $x \succ 1$) and conduct Update quarters between each Refresh. Daoust et al. (2008) provides a comprehensive review of the future work that lies ahead for the continuous improvement of QFS. In this paper, we have discussed the implementation of merely one of those review items.

ACKNOWLEDGEMENTS

The authors would like to sincerely thank Jean Leduc, Harold Mantel, Haig McCarrell and Zdenek Patak for their key insights which contributed to improving the quality of this paper.

REFERENCES

Statistics Canada. (2007). “North American Industry Classification System (NAICS) 2007 - Canada”. *http://www.statcan.gc.ca/subjects-sujets/standard-norme/naics-scian/2007/list-liste-eng.htm.*

Lavallée, P., and Hidiroglou, M. (1988). “On the Stratification of Skewed Populations”. *Survey Methodology*, Volume 14, Number 1, 33-43.

Kish, L. and Scott, A. (1971) “Retaining Units after Changing Strata and Probabilities”, *Journal of the American Statistical Association*, Volume 66, Number 335, 461-470.

Daoust, P., Bohossian, N., Gauthier, S., Demers, G., and Nicholas, J. (2008). “Quarterly Survey of Financial Statistics for Enterprises”. *Proceedings of the Survey Methods Section, Statistical Society of Canada.*