

THE REDESIGN OF THE CANADIAN QUARTERLY RETAIL COMMODITY SURVEY

Catherine Dufour and H el ene B erard, Statistics Canada

Catherine Dufour, Business Survey Methods Division, Statistics Canada, Ottawa, K1A 0T6, balacat@statcan.ca

KEY WORDS : Commodity; Multivariate allocation; Redesign; Retail trade; Two-phase design.

Abstract

The Canadian Quarterly Retail Commodity Survey (QRCS), developed in 1997, collects data and produces estimates of the sales distribution for retail commodities such as food, clothing and furniture. The survey uses a two-phase stratified sample design, where the first phase is the Monthly Retail Trade Survey (MRTS). Both surveys are now undergoing a redesign, in part to provide estimates for the new North American Industry Classification System (NAICS), and to take full advantage of administrative data from the Good and Services Tax (GST) program. The new QRCS will retain its two-phase stratified sample design, with a multivariate optimal sample allocation to maximize efficiency of the design for major commodities. Nevertheless, many changes will be brought to the design. For example, to reduce respondent burden on smaller units, the survey will not collect data for a large portion of its population; these units however represent less than 5% of total sales. The production of commodity estimates for the non-survey and survey population, with the added constraint that total sales must be benchmarked to those of the MRTS, constitutes an interesting challenge. This paper discusses the changes brought about by the redesign and presents the survey processes that will be implemented in the new QRCS survey such as sample design, sample and frame maintenance, edit and imputation, and estimation methodology.

1. Description of the current QRCS

In January 1997, Statistics Canada launched the Quarterly Retail Commodity Survey, whose mandate was to collect detailed information about retail commodity sales in Canada. Prior to the introduction of the QRCS, Statistics Canada identified the lack of detailed information on commodities as a data gap. In particular, this information was required by the System of National Accounts. A two-phase sampling plan was developed to take advantage of information gathered by the Monthly Retail Trade Survey (MRTS), which measures total retail sales in Canada. This allowed the QRCS to be developed more economically and within a tighter time frame. From a statistical perspective, first-

phase information created efficiency gains in a number of areas, including the selection of the same respondents, the use of existing systems and staff, the restratification of the first-phase sample before selection of the second phase, and the use of auxiliary information in sample allocation, edit and imputation and estimation. The QRCS sample consists of 10,500 companies in Canada. Information on the sales of 117 commodities is collected each month. These commodities are assembled into nine major commodity groups such as food, clothing and accessories, and furniture and appliances. The estimates are produced and published on a quarterly basis. More details can be found in Brodeur et al. (1999).

The MRTS and the QRCS are undergoing a complete redesign. The newly redesigned MRTS and QRCS are projected to be in production in 2003 after a five-month parallel run with both the old and new designs. This paper describes the reasons for this redesign. The features of the new QRCS are also described and the major differences compared to the current design are explained.

2. Redesign of the MRTS and QRCS

The current MRTS and QRCS use a stratified sample design to which many improvements have been made, strengthening data quality. However, a redesign was needed to address several main issues, such as reducing cost and respondent burden, updating computer systems, and harmonizing various concepts with those used by other annual surveys that have also been redesigned. Because of the two-phase design, changes to the first-phase sample (MRTS sample) affect the second-phase sample.

The new MRTS will continue to use a sample design stratified by industry, geography, and size with simple random sample selection in each stratum. As well, a multivariate sample allocation will be kept for QRCS. However, many changes will be brought to different aspects of the sampling designs of MRTS and QRCS. Specifically, the redesign includes the conversion to the new North American Industry Classification System (NAICS), the inclusion of non-employers, the introduction of a non-surveyed portion, a new sampling unit and new size measures and stratification variables (see B erard, 2001).

2.1 Populations

The MRTS and the QRCS will continue to use the Business Register (BR) to extract their frame. The BR is a list of Canadian businesses maintained by Statistics Canada. Businesses on the BR are represented by a hierarchical structure with four levels, with the enterprise at the top, followed by the company, the establishment and the location. An enterprise can be linked to one or more companies; furthermore a company can be linked to one or more establishments, and an establishment to one or more locations.

The target population for the new MRTS and new QRCS includes Canadian retailers that have establishments coded under NAICS. The MRTS and QRCS will again exclude certain sub-sectors such as electronic shopping, mail-order houses, vending machine operators and direct selling establishments. The conversion to NAICS will bring in a new sub-sector, computer and software stores, which was not included in the current designs based on the Standard Industry Classification of 1980 (SIC80).

The survey population for the current MRTS and QRCS includes employers only, one reason being because the variable used for size measure is unavailable for non-employers. To take into account the non-employer portion, an adjustment is performed at the estimation stage. The survey population for the redesigned MRTS and QRCS includes both employers and non-employers. All establishments coded as retail under NAICS, except the excluded sub-sectors, are included in the survey population.

One of the main impacts on the designs is the introduction of a non-survey portion due to the application of exclusion thresholds to reduce respondent burden on smaller units. These thresholds are applied at the cell level (combination of trade group and geography). Sampling units falling under the size thresholds are excluded, but account for less than 5% of total sales at the cell level. Analysis has shown that after the exclusions, the retail population is reduced from 211,000 establishments to 129,000, thus approximately 39% of all establishments are excluded. Estimates will be produced for the non-survey portion for both MRTS and QRCS through the use of administrative data.

2.2 New Sampling Unit

For the redesign, the sampling unit will no longer be the company, but the cluster of establishments, defined as establishments of the same enterprise that operate in the same industry group and same geography. Both MRTS

and QRCS use this new sampling unit, which is also now used in the Annual Retail Trade Survey (ARTS).

2.3 New Size Measure

A new size measure has been developed for the redesign. It is created by using a combination of independent survey data and three administrative variables: the Gross Business Income (GBI), the T2-revenue (the annual revenue from the income tax reports of corporations), and the sales reported through the Goods and Services Tax program (GST sales). The independent survey data consist of the annual sales available from respondents to the current MRTS and the ARTS. The samples from these two surveys are both independent of the redesigned MRTS. For sampling units to the redesign that were also respondents to the current MRTS or the ARTS, the size measure will be set to the annual sales obtained from the most up to date information from these surveys. Recent respondents' data are deemed to be a better indicator of size than administrative data. When respondents' data from independent surveys are not available, the size measure is set equal to the largest of the administrative variables among the GBI, the GST sales and T2-revenue that are available for that business. All three administrative variables are available for employers. For non-employers, the GBI is not available; only the GST sales and the T2-revenue may be present. Note that employers represent 92% of retail total sales. A 10% reduction in the misclassification rate of employers is observed when the size measure relies on the three administrative variables instead of the GBI alone. This new size measure is particularly efficient in identifying large businesses.

2.4 First phase stratification and sample allocation (MRTS)

In the current MRTS design, the GBI for employers was used as the size measure to stratify within each geographic region and industrial group. Each stratum (trade group and geographic combination) is divided by size into three substrata: a take-all stratum and two take-some strata, one containing medium-sized companies and the other small companies. The take-all strata include all companies with a complex structure, that is, companies that operate in more than one trade group or geographic area, as well as companies with a GBI above a certain limit. The thresholds were calculated using a method developed by Hidioglou (1986). The sample was allocated in each take-some stratum proportionally to the square root of the GBI.

In the redesign, the stratum boundaries are identified by the Lavallée-Hidiroglou algorithm (see Lavallée-Hidiroglou (1988)). This algorithm allows for an optimal allocation of the sample based on target coefficients of variation, and the creation of one take-all stratum and as many take-some strata as desired. The new sampling design includes one take-all stratum and at most two take-some strata. The sample is allocated in each take-some stratum proportionally to the square root of the new size measure. The target coefficient of variation (CV) of sales will remain at 1.2% at the national level, 2.5% at the geographic level and 3.5% at the trade group level.

2.5 Second phase stratification and sample allocation

In 1997, the first-phase sample was restratified prior to selection of the second-phase sample. This will not occur under the redesign, the two surveys will use the same stratification. This is simpler operationally and eliminates a gap in the data, in that collection can begin immediately for the QRCS along with the MRTS collection. As well, all MRTS take-all units will become QRCS take-all units. There will be sub-sampling only in the take-some strata. Again, this is an operational simplification, and reduces the likelihood of having large units in the take-some strata with large weights.

The QRCS is intended to provide sales estimates for many commodity groups, requiring a multivariate sample allocation. The 1997 design used an adaptation of Bethel's (1989) algorithm since there was no conventional solution to the problem of optimal multivariate allocation when using a stratified two-phase sampling design. For the redesign, an automated allocation procedure was developed which also uses the Bethel Algorithm. Since the multivariate allocation relies on commodity information for the nine major groups under the new design, population commodity estimates and measures of variability for all strata had to be estimated using the current survey. Target CV's are set at 7.5% for major commodities within each trade group. An oversampling procedure is applied after the initial allocation to take into account non-response, misclassification, deaths, minimum sample sizes in the take-some strata and maximum weights per strata. The sample size under the redesign is roughly 10,000 clusters of establishments, which translate to approximately 7,500 Collection Entities (CE's). A CE is defined as all sampled establishments in the same trade group within the same statistical company. This compares to 10,500 companies and 8,500 CE's in the current design.

2.6 Samples and Frame Update

For the first production month, the population will be stratified according to industry, geography, and the new size measure, and then the initial first phase and second phase samples will be selected. For the following months, the samples will be updated to reflect the changes in the population. Different procedures will take place for births, deaths and misclassified units. Births and deaths will be defined at the sampling unit level, that is, the cluster of establishments within an enterprise.

Births (new clusters of establishments) are identified every month via the latest BR universe. They will be stratified according to the same criteria as the initial population. A sample of these births will be selected according to the sampling fraction of the first and second phase strata to which they belong. This sample of births will be added to the monthly samples.

Deaths occur on a monthly basis. The status of these businesses is updated on the BR by using administrative sources and survey feedback, including feedback from the MRTS or QRCS. A death can either be a cluster of establishments which have ceased their activities (out-of-business), or whose major activities are no longer in retail trade. Deaths identified by MRTS or QRCS will remain in the sample but will not be sent a questionnaire. The business will continue to contribute to the level and variance estimates with a value of zero for its sales. Some dead units will be removed twice a year from the frame and the sample in an unbiased manner as described in Trépanier et al. (1998). These steps are necessary since MRTS and QRCS are non-independent sources of information that contribute to update the BR.

Due to the dynamic nature of retailers, it is expected that over time, many businesses will change size and thus belong to the wrong size stratum. Units that are misclassified may cause undesirable fluctuations in the estimates. To deal with this problem, three types of restratification exercises are being considered for the redesign; monthly, annual, and every five years. Details can be found in Majkowski (2001).

3. Collection

In the new surveys, the collection strategy will be similar to what is currently done. The units of collection are the same for MRTS and QRCS and data collection is combined for the two surveys, using the same infrastructure. Although the surveys have different questionnaires, collection and follow-up are done for

both surveys through one monthly contact. The MRTS gathers total monthly retail sales, while the QRCS asks respondents for a breakdown of those sales by commodity groups. The QRCS questionnaire lists 117 different commodities, which are further regrouped into 27 commodity groups, or combinations of parts and totals. The respondent is asked to report sales for all 144 of these levels, with the total sales for all major groups equal to the total retail sales from MRTS. The respondent can report its sales by commodity group as either a dollar amount or as a percentage of its total retail sales; if unable to provide this data, the respondent is asked the types of commodities sold. This information is used at the edit and imputation stage to determine which fields need to be imputed.

The majority of companies are in the survey sample for extended periods, thus questionnaires are tailored to the commodities reported in previous responses. The first time a unit is contacted for the QRCS, a profile is created listing the commodities usually sold. The profile is initially used in preparing the tailored questionnaire and later in edit and imputation, and is updated regularly. The tailored questionnaire eases response burden and helps boost the response rate. The respondent may report data on a monthly, quarterly, semi-annual or annual basis. Annual reporting may be done when the distribution of sales does not vary throughout the year; data are imputed for months they are not reporting.

With the current survey, initial contact is made through mail-out, with respondents having the option to report subsequently by mail, telephone, or other methods. At the present time 60% of collection is done by telephone, with the remainder done mostly by mail. For the redesign, initial contact will be made by telephone, and it is foreseen that most collection will be done by telephone as well.

4. Edit and Imputation

An overall strategy similar to that of the current survey will be used for the edit and imputation processes of the new QRCS. The main impact of the redesign on edit and imputation is due to the change in industrial classification by the introduction of NAICS, resulting in new trade groups. For example, all the edit rules linking commodities to industry group have to be reviewed. As well, the imputation groups or data groups created under the SIC80 classifications have to be reanalyzed under NAICS.

The edit and imputation process for QRCS can sometimes be complex. This is due to the large number of different commodities, some of which vary from

company to company even within the same trade group, and to the many totals and subtotals of commodity groups (see Bérard et al. (1999)). The editing and imputation system consists of three main modules: pre-editing, automated editing, and imputation.

4.1 Pre-Editing

The pre-editing module performs a series of verifications on data supplied by units that contribute most to the estimate of total sales in each retail trade sector. Such units may be either large businesses or small businesses that have a high sampling weight. The data they provide are verified to ensure that sums of parts and totals add up, that reported commodities match the type of business, and that there are no drastic changes in sales from month to month or year to year. Subject matter experts examine any data that fail pre-editing, where the most obvious errors are corrected, or respondents are contacted for clarification.

4.2 Automated Editing

All data, including those which have undergone pre-editing, must go through the next stage, automated editing. The object of automated editing is to identify fields requiring imputation, while altering as little as possible the data reported by respondents. Automated editing finds erroneous data that must be replaced with imputed values when there is inconsistency between parts and totals. Verifying sums of parts and totals can be quite complex because subtotals are added together to form other totals. Adjusting one of the subtotals to resolve an inconsistency with the sum of its parts may induce an inconsistency in the sum of the subtotals with the grand total.

When encountering total nonresponse, the automated editing system determines which of the fields involved should be imputed to zero and which should be imputed to a positive value. The profile created during data collection, historical data (for the previous month or the same month of the previous year) and even the unit's industrial classification are used in this process. Prior to actual imputation, additional verification is performed on records that might be used to form imputation groups, as described in the next section, to ensure that only homogenous units will be used. For example, a women's clothing store that also sells cosmetics would not be used for imputation; if so, it might generate cosmetic sales for all non-respondent women's clothing stores.

4.3 Imputation

The first step in the imputation process involves

defining imputation groups, or sets of homogeneous units. A value imputed to a unit will be derived from the values of respondents belonging to the same imputation group. The QRCS imputation groups are defined on the basis of the most recent information about the unit's industrial classification, geography and size. Different imputation groupings were defined, some at a less detailed level. For example, an imputation group can be defined at a small geographic level or at a broader level. When imputation is not feasible at a more detailed level, for example if there is an insufficient number of respondents to calculate trends, then imputation is performed at the next less-detailed level.

From the outset, QRCS used ratio imputation and adjusted historical imputation. Although edit and imputation are applied to the dollar values of commodity sales, the survey is much more concerned with the distribution of commodities, that is, the proportion of sales of each commodity in relation to total retail sales. For this reason ratio imputation methods were initially preferred over adjusted historical methods, but were found to produce inconsistent results. Thus, the QRCS currently uses adjusted historical methods. For total non-response, data from the same month of the previous year for that unit is used; if unavailable, then the previous month is used; and again if unavailable, donor imputation within imputation groups is used. For partial non-response the imputation groups are used exclusively. Only one commodity is imputed at a time, although block imputation, where donors satisfy a group of related variables such as parts of a given total, is being considered for the redesign. Finally, since imputation does not ensure that the parts will add up to the totals, it is followed by a prorating step, where QRCS total sales are prorated to MRTS sales at the questionnaire level.

5. Estimation

The goal of the QRCS is to produce estimates for the distribution of total retail sales among various commodities. The source of the level of these sales is the MRTS. Total sales for the redesigned MRTS will include total sales from the survey portion and total sales from the non-survey portion. Sales from the non-survey portion will be modelled through the use of GST sales. QRCS sales obtained from the survey portion will be benchmarked at the trade group level to total MRTS sales, which include sales of both the surveyed and non-surveyed portions as described earlier.

Estimation is done through the Generalized Estimation System (GES), an automated estimation software developed by Statistics Canada. Variances are calculated using a formula developed with the Taylor

linearization method and adapted for a two-phase stratified design (see Binder (1996)). Because the collection and estimation methodology are performed on a monthly basis, but results for the QRCS are published quarterly, co-variances also have to be calculated for the three months of each quarter.

6. Future Work

A five-month parallel run is scheduled to start in 2003. During this time both the current and redesigned surveys will be in production, collecting data and producing estimates. During the parallel run period, results from both surveys will be monitored and compared, and if necessary, adjustments will be made to the different survey processes.

References

- Bérard, H. (2001). The Redesign of the Monthly Wholesale and Retail Trade Survey of Statistics Canada, *Proceedings of the Survey Methods Section*, Statistical Society of Canada, 81-86.
- Bérard, H., Brodeur, M. and St. Pierre, M. (1999). The Retail Commodity Survey, *Proceedings of the Survey Methods Section*, Statistical Society of Canada, 141-146.
- Bethel, J.W. (1989). Sample Allocation in Multivariate Surveys, *Survey Methodology*, **15**, 47-57.
- Binder, D.A. (1996). Frequency Valid Multiple Imputation for Surveys with a Complex Design, *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 281-286.
- Brodeur, M., Bérard, H. and St. Pierre, M. (1999). The Canadian Retail Commodity Survey, *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 163-168.
- Hidiroglou, M.A. (1986). The construction of a Self-Representing Stratum of Large Units in Survey Design, *The American Statistician*, **40**, 27-31.
- Lavallée, P. and Hidiroglou, M.A. (1988). On the Stratification of Skewed Population, *Survey Methodology*, **14**, 33-45.

Majkowski, M. (2001). Maintaining Estimate Quality and Easing Response Burden in a Subannual Business Survey, *Proceedings of the Business and Economic Statistics Section*, American Statistical Association, to appear.

Trépanier, J., Babyak, C., Marchand, I., Bissonnette, J. and St-Pierre, M. (1998). Enhancements to the Canadian Monthly Wholesale and Retail Trade Survey, *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 487-492.