Sampling in the 2015 Bank of Canada Retailer Cost of Payments Survey

Angelika Welte[1]
October 7, 2016

Abstract
The Bank of Canada undertook the large scale Retailer Survey on the Cost of Payment Methods in 2015 (2015 RCPM). This paper describes and discusses the sampling methodology used in this survey with a focus on the challenges of voluntary business surveys. Recommendations for sampling strategies in future retailer surveys are offered.

Key Words: Survey design, sampling frame, business survey.

1. Survey context

The Bank of Canada as the sole issuer of banknotes in Canada has a policy and research interest in the use of payment methods by Canadian retailers, consumers and other stakeholders in the retail payment system. Smooth and efficient retail payments depend to a large part on the costs borne by each stakeholder and on the fees paid from one stakeholder to the other. Therefore, in 2014, the Bank of Canada's Currency department initiated a large scale research project, the Cost of Payments Study, to collect cost data on points-of-sale (POS) transactions from retailers, financial institutions, cash-in-transit companies and consumers. The Bank of Canada followed with this study the example of other central banks, such as thirteen countries in the European Union (Schmiedel, Kostova, and Ruttenberg (2013)) and the Reserve Bank of Australia (Stewart, Chan, Ossolinski, Halperin, and Ryan (2014)), and public authorities such as the European Commission (ECDGC (2015)). The Bank of Canada Cost of Payments Study, as many of the other studies, focused on cash and payment card transactions at a physical POS where a consumer purchases a good or service from a business. It also collected some data on cheques, but not on online transactions, credit transfers and direct debits. Business-to-business transactions are likewise out of its scope. While the findings of the Cost of Payments Study will be published as Bank of Canada staff discussion paper in 2017, this report focuses on the survey methodology for the data collection among Canadian retailers, termed the *2015 Retailer Survey on the Cost of Payment Methods (2015 RCPM)*. It first summarizes challenges encountered in an earlier payment-focused Bank of Canada Retailer Study conducted in 2006 and then highlights the methodological changes implemented for the 2015 RCPM, followed by two technical sections on the 2015 RCPM sample frame and sampling procedure. Key issues with the 2015 RCPM sampling methodology are discussed before a brief conclusion is reached.

2. Methodological revision of 2006 Retailer Study for 2015 RCPM

For the 2006 (Bank of Canada) Retailer Study, the precursor of the 2015 RCPM, a marketing research firm had been commissioned to collect data from retailers through Computer Assisted Telephone Interviews (CATI). The development of the survey methodology, including the choice of a sample frame and the weighting of the final responses, were carried out by the same firm. To reach a representative sample, the marketing research firm made

[1]Bank of Canada, 234 Laurier Avenue West, Ottawa, Ontario K1A 0G9, Canada (E-mail: awelte@bankofcanada.ca).

phone calls until the quotas in the contract were reached. Quotas were set for region, industry and business size. The survey had a response rate of 5 per cent among dialed numbers and a respondent sample size of 500. Some quotas were not reached and Arango and Taylor (2008a) caution against generalization of the survey's findings due to a high margin error. In Arango and Taylor (2008b), they based cost calculations on an even smaller sample of 35 respondents from a follow-up paper survey. The follow-up survey was necessary since key questions on the cost of payments had suffered from high item nonresponse in the CATI survey. The 500 responses also came from a mix of chain and independent stores, but were not adjusted for repetition of the same chain within the sample.

Three main concerns for the 2015 RCPM emerged from the 2006 Retailer Survey: First that, due to high response burden and the voluntary nature of the study, overall response rates in 2015 would be low and that the completed questionnaires would suffer from item non-response. Second, that the collected responses would not constitute a representative sample of Canadian retailers. Third, that the Bank of Canada team sought greater control of the sampling procedure and the inclusion probability of each business. The methodological choices for the 2015 RCPM aim at improving these three issues.

To boost response rates and reduce respondent burden, the survey team relied on the Tailored Design Survey Method (Dillman, Smyth, and Christian (2008)) and feedback collected from business owners during the testing of the questionnaires (sections 2.1 to 2.3). For the second concern, stratification was employed in combination with responsive sampling design (sections 2.4 and 4) so that survey effort would increase where low response rates or large margins of error were encountered. To gain control over sampling, the Bank of Canada constructed a survey frame, sampled all businesses in-house and then provided the inclusion probabilities where possible (sections 2.3, 3 and 4). The majority of the data collection and processing was still carried out by a marketing research firm as the Bank of Canada did not have the resources for those tasks.

2.1 Survey mode

The 2015 RCPM was planned as a mixed-mode data collection survey, meaning that survey responses would be submitted through several channels: paper questionnaires, online questionnaires and telephone interviews (CATI). While Dillman (2006) identified survey mode as a source of instability in household surveys, flexible survey modes can also reduce respondent burden and increase response rates. Cognitive testing of the questionnaire with a small number of businesses in early 2015 confirmed that completion by phone would take too much time during a typical work day since the questionnaire consisted of eight pages and required respondents to look up details in their financial records. The majority of the RCPM sample therefore received paper questionnaires by mail. The personalized online questionnaire was made available to every other sampled business (selected at random) and all sampled businesses were given online access on reminder postcard several weeks later. Unpersonalized online access was also available on the Bank of Canada's website. Lastly, phone calls were used for nonresponse follow-up and to boost sample size in certain strata.

2.2 Incentives

Incentives offered to respondents in the 2006 Retailer Study were charitable donations in the name of the responding business and a special copy of the study report. Similarly, the 2015 RCPM also offered incentives to respondents since the Bank of Canada felt that incentives not tied to responding may be viewed as an inappropriate use of funds although Dillman, Smyth, and Christian (2008) demonstrate that advance cash incentives effectively increase response rates. Advance letters signed by the Governor of the Bank of Canada

and letters accompanying the survey package signed by the Chief of the Currency Department explained that the data collected in the study would help gain insight into the cost of payments and emphasized that it would only be used for Bank of Canada research. Chen, Henry, Huynh, Shen, and Vincent (2016) report that a similar letter had improved participation in a Bank of Canada consumer survey. Besides the letter, the questionnaire allowed the business to select any number of the following incentives: (1) the final study report, (2) a detailed study report by industry, region and size, (3) a webinar presentation of the study report, (4) a certificate of appreciation, and (5) entering into a draw for a tablet computer. The draw for a tablet computer turned out to be the most popular incentives, but many respondents also requested a study report.

2.3 Sampling frame and survey instruments

The survey frame was mainly based on over 400,000 downloaded business units from the Dun&Bradstreet (D&B) database which were combined with information on the largest retail and restaurant chains in Canada. D&B has been used extensively by the Bank of Canada for other business surveys (de Munnik, Dupuis, and Illing (2013)) and also by the Federal Reserve Bank of San Francisco for the Pilots of their Cash Payments Survey. It lists an address, employee counts and industry information in the form of the North American Industry Classification System code (NAICS) for the majority of units, hence facilitating stratified sampling. In business surveys, the economically most significant firms are usually included in the sample with probability one, forming a take-all (TA) or certainty stratum. Bank of Canada researchers combined information from D&B, Restaurants Canada[2], the Monthly Retail Trade Survey (Statistics Canada (2014)) and the Retail Council of Canada for the definition of the TA stratum.

Choosing the correct survey unit is an important part of business survey methodology (Rivière (2002)). In the 2015 RCPM, the contacted survey unit had to be capable of providing data on payments while also being authorized to release this information. To avoid duplication, the unit should also be at the highest level in its organizational hierarchy where data on payments are available. Most businesses in Canada are simple and consist of just one unit, while a small fraction of businesses are organized into a complex hierarchy of multiple units. Complex businesses however contribute a significant portion of economic activity (Statistics Canada (2010)). Stratification by firm structure was therefore recommended. The Single Location stratum (SL) were independently owned and operated, mainly small and medium sized, businesses that do not operate under a brand name or banner of a large chain. The HQ stratum consisted of clusters representing large chains, multi-unit businesses with a complex structure and potentially several locations, businesses falling under the same brand and other businesses with large assets or revenue. Since an HQ may represent several units in the original database, we also refer to the units in the HQ stratum as "chains" or "complex businesses." The aforementioned TA stratum was included in the HQ stratum as the HQ TA stratum. Section 3.1 gives a detailed technical description of the strata and cluster construction.

For SLs on the frame the survey and the responding unit were identical. For clusters, the survey unit was the cluster. Since the head office may not be able to report payment activities such as the time spent every day on counting coins and banknotes, responses were also required from individual locations. As suggested by Dillman, Smyth, and Christian (2008), the survey instruments are tailored to the business structure:

1. Single location (SL) questionnaire for businesses that are independently owned and

[2] http://foodserviceandhospitality.com/june-2014-digital-issue/

operated, do not operate as part of a chain or banner, and are not classified as headquarters.

2. The questionnaire for HQ clusters is divided into two sections:

   (a) The Head quarter questionnaire (HQ) for the head offices,

   (b) Branch or location questionnaire for locations, branches or franchises.

   The head office needed to coordinate and authorize completion of all questionnaire sections.

2.4 Stratification

The units on the sampling frame were first stratified according to their *structure* and then according to *region, industry and size* as routinely employed in business surveys (Table 1). The HQ TA stratum contains the largest retail chains in Canada, the HQ Take Some (TS) stratum other large and complex businesses, and finally the Single Location (SL) stratum the remaining independently owned and operated businesses. SL has the largest number of units, but each of them contributes a small fraction of POS transactions, while HQ TA has a small number of units with much more POS transactions.

Results from the 2013 Methods-of-Payment Survey (Henry, Huynh, and Shen (2015)) indicate that region, industry and size may be correlated with acceptance of payment methods. Stratification by all three dimensions -region, size and industry- was used for the SL sample, the HQ TS sample was stratified along region and industry and the HQ TA stratum by industry. The separation of HQ TA and TS already takes care of size differences between HQs. The regional strata were Atlantic (AT), Quebec (QC), Ontario (ON), Prairies (PR) and British Columbia (BC). For the size of the businesses, Stratum A were single locations with less than 5 employees or a missing number of employees, Stratum B those with at least 5, but less than 50 employees, and Stratum C those with at least 50 employees. Industry strata were given by the 2-digit NAICS (44-45[retail trade], 72 [food service and drinking places], and 81 [repair and maintenance, personal and laundry services]).

	SL	HQ TS	HQ TA
Revenue per unit	Low	$\rightarrow$	High
Number of businesses	High	$\rightarrow$	Few
Stratification	R, N, S	R, N	N

Table 1: Stratification levels: R(region), S(size), N(industry).

3. Sampling frame construction

A customized sample frame was constructed for the 2015 RCPM to accommodate the desired stratification by firm structure (Figure 1).

3.1 Construction of clusters for HQ frames

The construction of the clusters for the HQ TA frames was based on a fixed list of national chains that the HQ TA frame had to cover, while there was no such list for the HQ TS frame. For the HQ TS frame, the clusters were formed using relationships between units derived from D&B.

3.1.1 HQ TA

The HQ TA frame consists of approximately 100 restaurants chains (NAICS 72) and 80 retailers (NAICS 44-45) in Canada. Services (NAICS 81) are not included in the HQ TA due to lack of access to a reliable list of the largest providers of consumer services in Canada at the time. The technical implementation of the clustering procedures takes the following steps:

1. Split the D&B database into three databases: retail (NAICS 44-45), restaurants (NAICS 72), and services (NAICS 81).
2. For NAICS 44-45 and NAICS 72
   (a) Standardize the spelling of the following D&B variables: companyname, doingbusinessas, immediateparent, ultimateparent.
   (b) Form clusters by matching company names, parent names and operating names to the list of largest businesses in the NAICS stratum.
3. Eliminate all units within a cluster from the data base and proceed to the HQ TS

3.1.2 HQ TS

In the HQ TA step, units associated with the largest Canadian retail and restaurant chains are filtered out of the database. The HQ TS step accounts for other large and complex enterprises. Clusters are formed based on:

1. Classification as "headquarter" in D&B,
2. Revenue exceeding 15 million $US dollars,
3. Total assets exceeding 10 $US million dollars,
4. Same immediate or ultimate parent within a NAICS,
5. Repeated names within a NAICS.

3.1.3 SL

In the construction of the HQ cluster, many units that are "single locations" in D&B were sorted into clusters. To avoid duplication, the SL frame consists of all "single location" units D&B that are not contained in any HQ cluster. The SL CATI frame further consists of the SL frame units listing a telephone number in D&B.

4. Sampling procedures and design weights

Sampling and data collection proceeded in two phases, where the first phase was divided into three waves. The survey process is schematically represented in Figure 1. The activities on the left side were performed internally while those on the right were outsourced to the marketing research firm. Bank of Canada researchers were involved at every stage of the sample selection and kept in continuous contact with the data collection firm who performed address verification, mail-outs, and phone calls. The Bank of Canada monitored the margin of error for two benchmark variables from the SL responses, namely amount of

cash held on the premises and the value of cash transactions. If the margin of error was too high, additional responses had to be obtained were required.

In parallel to the description of the sampling procedure, this section also explains the approximation of inclusion probabilities π_i for the units i on the SL frame.

4.1 Cluster sampling of HQs

The sampling of HQs was a 3-step process. In the *first step*, a cluster of businesses was selected. In the *second step*, a contact unit was selected within the cluster. In case the Bank of Canada had a personal contact with the head office of the chain, this person was contacted. In case not, the contact unit was chosen from the D&B units in the cluster: If the cluster contains units of locationtype "headquarter" in D&B, the "headquarter" with the largest revenue was the contact unit. Otherwise, the unit with largest revenue was the contact unit. In case of a tie, one of the top revenue units was sampled at random. For large chains, the Bank of Canada's contacts gave better response rates than mailing a package to the largest revenue unit in D&B.

In the *third step*, the contact unit selected, at their convenience, up to three branches or locations. Responses are then obtained from the contact units and from the branches (locations). The inclusion probabilities for the HQ sample could not be calculated since convenience (or familiarity) sampling through personal contacts was the main recruitment method for HQs, see also Jiongo (2016).

4.2 Phase 1

Phase 1 for SLs consisted of three waves for which samples were drawn at random within cells h defined by three stratification variables, namely region, size and industry.

Wave 1 and 2 initial draw The 2006 Retailer Survey served as basis for sample size calculations. The targeted number of responses from SLs in each cell for Phase 1 was based on a set level of precision when survey weights were obtained from the raking procedures described in Deming and Stephan (1940). Furthermore, the minimum number of invitations for each stratum was 2,400 and additional invitations were allocated proportionally to the size of stratum on the SL frame.

Wave 1 and 2 replacement draw The sample size for the addresses replacing invalid addresses from the initial draw was calculated from the number of required replacements and the ratio of valid to invalid addresses so that the Bank of Canada staff could assume with 95 percent confidence that address screening of the replacement addresses would leave enough addresses to replace the invalid addresses in each cell h.

Wave 3 CATI booster sample draw Analysis of the responses collected in Wave 1 and 2 determined the need for additional responses since the margins of error were too high for the benchmark variables in certain cells. In these cells, the Wave 3 sample was drawn among units with a phonenumber.

4.2.1 Phase 1 Inclusion Probabilities

Inclusion probabilities for each wave and draw W are obtained by dividing the number of sampled entries n_h^W in a cell h by the frame size N_h^W, $\pi_i^W = \frac{n_h^W}{N_h^W}$ for all i in h. If a unit i was not on the frame for a wave W, then $\pi_i^W = 0$. In particular, for Wave 3,

the frame consisted of units with a phone number in the contact information and hence $\pi_i^{W3} = 0$ if i does not have a phone number. The thus obtained inclusion probabilities are $\pi_i^{W1}, \pi_i^{W2}, \pi_i^{W2r}$ and π_i^{W3}.

Recall that due to frame revisions, the frame sizes vary across waves (see the discussion section 5.1) Since a unit is included in at most one sample, an approximation of the Phase 1 inclusion probabilities is given by the sum of the four inclusion probabilities calculated so far

$$\pi_i^{P_1} = \pi_i^{W_1} + \pi_i^{W_2} + \pi_i^{W_{2r}} + \pi_i^{W_3}.$$

Finally, the Wave 2 sample also contains businesses recruited through personal contacts. No attempt has been made to calculate the inclusion probabilities.

4.3 Phase 2

Phase 2 aimed at filling the quotas in cells that are still considered underfilled after the booster sampling. The desired number of responses for Phase 2 in a cell h was determined as for the CATI booster sample, based on the same benchmark variables and all responses from Phase 1. Since the Bank of Canada researchers had access to cell response rates from Phase 1, they were confident about the required sample sizes $n_h^{P_2}$ and drew the entire Phase 2 sample at the beginning. To avoid returned survey packages in Phase 2, the marketing research firm made screening phone calls before mailing out the survey package to validate the addresses and NAICS on the frame as well as that the business accepts cash, debit cards or credit cards for POS transactions.

4.3.1 Phase 2 inclusion probabilities

The sampling proceeded in two steps. In the first step, the address frame was partitioned by phone number. Since phone numbers do not uniquely identify businesses, one unit was sampled at random for each phone number. In the second step, the phone numbers sampled in the first step are stratified by region, NAICS and size. In each cell h, $n_h^{P_2}$ phone numbers were sampled at random from the telephone numbers. Phone numbers belonging to businesses sampled in Phase 1 were not eligible for the draw. Instead of the exact calculation of the inclusion probabilities, the following approximation is used: Denote by r_t the number of repetitions of the phone number t on the address frame. The probability to include a business with phone number t on the phone frame is $1/r_t$ since one business with that phone number is picked at random. The cell weight $w_{t,h}$ of the phone number t is the fraction of businesses with phone number t in cell h. The frame size for cell h in the second stage, $N_h^{P_2}$, is a random variable with expected value

$$E(N_h^{P_2}) = \sum_t w_{t,h}. \tag{1}$$

where the sum is over all phone numbers t. The probability for including a business i with phone number t from cell h in Phase 2 is approximated as

$$\pi_i^{P2} = \frac{n_h^{P_2}}{E(N_h^{P_2})}.$$

Businesses without phone contact on the frame have an inclusion probability of zero for Phase 2. Table 2 compares the approximated inclusion probabilities with the sampled fractions (for the cells targeted in Phase 2).

The calculation of the exact Phase 2 inclusion probabilities would be computationally intensive and would require enumeration of all possible samples.

Repetitions	Share of pairwise different phone numbers	Observed Inclusion in perc.	Approximated Inclusion in perc.
1	92.57	4.47	4.46
2	6.92	2.58	2.62
3	0.45	1.81	1.77
≥ 4	0.06	1.20	2.07

Table 2: Observed vs. Approximated Inclusion Probabilities in percent. Calculated on the set of businesses having a phone number and falling into the targeted cells for Phase 2.

For illustration, let the frame for Phase 1 be $\{1, \bar{1}, 2, \bar{2}, \bar{\bar{2}}, 3\}$. Elements $i, \bar{i}$ and $\bar{\bar{i}}$ have the same phone number. Assume that $n^{P_1} = 2$ and $n^{P_2} = 1$. It can then be shown by enumerating all possible samples, that the probability to include the element 3 in either the Phase 1 or Phase 2 sample is $\pi_3 = 13/15$. With the approximation, $\pi_3^{P_1} = 1/3$ and $\pi_3^{P_2} = 1/3$, so that $\pi_3^{P_1} + \pi_3^{P_2} = 2/3 \neq 13/15$.

4.4 Design Weight for SL

The overall inclusion probabilities are approximated as the inclusion probabilities for Phase 1 and Phase 2, truncated at 1:

$$\pi_i = \min(\pi_i^{P1} + \pi_i^{P2}, 1). \tag{2}$$

Note that a unit was included in at most one wave, so that, in the special case where the frames are identical for each draw and random sampling is used in each wave, this formula is actually exact. The sampling and calculation of inclusion probabilities are hereby completely described. Remaining challenges of technical and non-technical nature will now be discussed in section 5.

5. Lessons learned from the 2015 RCPM

Compared to the 2006 Retailer Study, the 2015 RCPM saw a much greater involvement of Bank of Canada staff. Staff constructed the sampling frame and devised a complex sample design to collect a nationally representative sample and to reduce the margin of error. Due to their hands-on experience and access to the sample frame, they were also able to identify further areas for improvement. Sections 5.1 and 5.2 are directly related to the sampling process for SLs, sections 5.3 and 5.4 concern the quality of information on the frame and in the D&B database, section 5.5 discusses the construction of clusters for the HQ frame and finally section 5.6 touches on calibration.

5.1 Sampling design

The 2015 RCPM was designed to consist of two phases and each of which could have several waves. Sampling and data collection in later stages was to be guided by paradata from earlier stages such as call records from phone calls, returned-to-sender (RTS) survey packages and observations from data processing and coding. The presented sample design performed well in terms of flexibility since the survey team was able to adjust the sample frame and data collection protocol during the data collection. The survey team could also allocate greater survey effort to cells where they observed low response rates or high variability of the outcome variables. Additionally, waves randomized the time when

the business received the questionnaire and protected against weekday and holiday effects. The design is complex and exact inclusion probabilities cannot be calculated, however.

If frame revisions can be avoided, a future survey could follow a simplified survey design. The following proposed adaptive collection design has one phase of sampling and can accommodate several waves of data collection. Response rate estimates need not be known and only the required number of responses is collected. First, order the businesses randomly within each cell. Then businesses are selected according to the random order until the desired number of responses is reached. Because the order is random, the selection probability is simply the index of the last selected business divided by the total number of units in the stratum. In practice, the sampling would be done in batches, still following the indices. Follow-up also proceeds according to the indices, skipping those who have submitted satisfactory questionnaires. In such a design, data collection effort such as the call attempts or the value of the incentives can be adapted based on paradata acquired during the survey (e.g. subgroup response rates). For example, Beaumont, Bocci, and Haziza (2014) show how to increase survey quality given a fixed survey budget using adaptive data collection for a CATI survey. Alternatively, when response rates are well-understood for each stratum and frame revisions are unlikely, stratified simple random sampling with just one draw (wave) is the simplest option. While this option offers less control over the exact number of responses, the inclusion probabilities as well as estimation and inference procedures are well understood.

5.2 Distribution of design weights

The design weight d_i of unit i is the inverse of the inclusion probability $\frac{1}{\pi_i}$. If S is a probability sample from a finite population P of size N, then unbiased estimators of the population size and the sample size (Levy and Lemeshow (2008)) are given by

$$\widehat{N} = \sum_{s \in S} d_s, \qquad (3) \qquad\qquad \hat{n} = \sum_{j \in P} \pi_j. \qquad (4)$$

In particular, if the population is divided in strata and the design is stratified random sampling, then we obtain likewise unbiased estimators for the stratum size N_h and sample size n_h. The variability of the design weights within a cell h as seen in Table 6 is mostly due to the sampling from multiple overlapping frames: After the initial Phase 1 draw, improvements to the clustering procedures resulted in a new frame for the replacement draws. Next, all units without a phone number were dropped from the frame starting with Phase 1 Wave 3. Sampling from the phone frame is further modified for Phase 2 by switching to a 2-step sampling procedure to deal with phone number duplication. Lohr (2011) introduces the single frame adjustment for design weights in multiframe sampling and this adjustment was used in the 2015 RCPM as described in Section 4. This adjustment preserves relationships between survey variables, hence is *internally consistent*. If identity (4) holds in each cell, the design weights are said to be *externally consistent* with the population size in each cell on the frame. The ratios $\hat{N}$ and $\hat{n}$ to actual population and sample size, respectively, are displayed in Figure 2. The figures confirm Lohr (2011)'s statement that the single frame adjusted weight may be externally inconsistent even if the weights are consistent for each frame. While calibration is usually used to ensure external consistency of the combined weight, future surveys should avoid the frame revisions that increased the number of frames in the 2015 RCPM.

5.3 Coverage error and unit nonresponse

In the 2015 RCPM, it is difficult to distinguish between (frame) coverage error and nonresponse error[3]. In particular, businesses with invalid addresses or returned survey packages (RTS) are treated as nonresponders as it is unknown whether they are still operating (eligible or "alive") or not (ineligible or "dead") although the ineligible businesses should be excluded from response rate calculations and the estimation of the nonresponse bias. Invalid addresses and RTS are both indicators that the information in D&B, and hence on the survey frame, is outdated. An added concern are response rates of 3 percent for all sampled businesses and of about 4 percent for all contacted businesses. Bank of Canada staff identified a need for nonresponse analysis, even if response rates are at best a incomplete measure of nonresponse bias and the response rate among eligible units may be higher. The Bank of Canada staff efforts to study nonresponse broadly followed Lineback and Thompson (2010)'s guidelines. As pointed out there, a full nonresponse follow-up study can be costly and time consuming since units must be pursued until their reason for nonresponse is discovered and data is collected from eligible nonresponders. In the 2015 RCPM, follow-up efforts were greatest for units in the HQ TA population since these units contribute a large portion of POS transactions and hence are considered influential in the estimation of total payment related costs for the population. Hatko (2016) uses auxiliary information from D&B to compare respondents and nonrespondents and develops a nonresponse model based on the frame variables. Bank of Canada research staff has been undertaking several initiatives to determine the eligibility of SL units in the D&B database. According to their probabilistic models, SL units with missing fields in the database had a higher than average propensity for RTS. After units at high risk for RTS were called to verify their status, about 25 to 30 percent of them were deemed ineligible. Based on these findings, the Bank of Canada added in Phase 2 a screening phone call to verify address and status of the sampled SL units. Again, a number of units were found to be ineligible. Since RTS appears to be linked to frame quality, a follow-up with the database provider for D&B is recommended to gather information on the current status of the sampled units and the last time they were updated.

5.4 Stratum jumpers

Stratum jumpers are units who turn out to belong to a different stratum than assumed during sampling. Stratum jumping with regards to measures of size can lead to influential units when the variable of interest is correlated with firm size, such as the number and total value of transactions in the 2015 RCPM. Due to the asymmetric firm size distribution, strata of small firms (size A and B) are larger than strata of large firms (size C, HQ TA and TS), so that the former have larger design weights than the latter. Businesses are also more likely to increase in size than to decrease. Therefore, stratum jumpers will often have large design weights and report a high transaction number or value.

Bank of Canada staff found that approximately 25 percent of the reported employee counts do not fall into the categories defined by the employee number in D&B. For the vast majority of the stratum jumpers, the reported category is either one size up or down from the category on the frame (i.e. jumping between A and B or B and C) and these were deemed unlikely to cause problems. Among the small number of potentially problematic stratum jumpers from size A (small, large design weights) to C (large, small design weights), two main types of stratum jumpers emerged. For the first type, the reporting unit

[3] Unit nonresponse is the failure to obtain survey data from a sampled business and occurs after the sampling step. For the remainder of the section nonresponse shall mean unit nonresponse.

appears different from the unit on the frame and is not a true stratum jumper. For the second type, "true stratum jumpers", the revenue on the frame is large and consistent with the reported employee count (but not the employee count on the frame). Future surveys may want to use revenue in addition to employee counts for the definition of size strata.

5.5 Cluster sampling

Cluster sampling is usually used in the presence of large within-cluster heterogeneity when a few clusters contain enough units to represent the population of interest. In the 2015 RCPM, intra-cluster (intra-chain) heterogeneity is assumed to be small, because cost of payments varies less between locations of the same chain than between chains. Therefore, as many clusters as possible must be sampled, but only a few units within each cluster. Franchises that are financially independent businesses are included in the clusters of their brand, as are company-owned stores. The assumption of intra-cluster homogeneity in the 2015 RCPM implies that franchises and company stores should incur similar costs for accepting payments. This assumption has merit since franchises and company stores are usually expected to offer a uniform customer experience and thus accept the same payment methods. The franchisor may also offer a pre-negotiated agreement with a payment service provider to the franchisee and fees paid for card payments will be similar within the chain.

After a lower than expected response from the Phase 1 mail-out sample, the HQ sampling relied mostly on contacts between Bank of Canada staff and the firms and, consequently, inclusion probabilities for the HQ sample could not be obtained. Future research may employ simulation methods as in de Munnik, Dupuis, and Illing (2013).

5.6 Sampling frame and population frame

While design weights are always based on the sampling frame, calibration adjustment can also use auxiliary information from another frame, the *population frame*. If the sampling frame is not representative of the population, calibration to auxiliary information may ensure representativeness of the final estimates for the whole population. The Statistics Canada Business Register (BR) suggests itself as a population frame for the 2015 RCPM. The BR is the basis of official business statistics in Canada, so that calibrated sample characteristics of the 2015 RCPM would be in agreement with these official statistics. Statistics from the BR are generally considered as reliable and up-to-date, owing to monthly quality assurance and mandatory participation (Statistics Canada (2010)). A drawback is that the units on the RCPM survey frame may not coincide with the units on the BR; SLs on the survey frame are not defined in the same way as the "statistical locations" in the BR. The difference is partially explained by D&B's definition of a single location and partially by the clustering of certain single locations into chains (see Sections 3.1 and 5.5). The BR also defines stratification variables in a different manner. For example, almost all headquarters and single locations in D&B have a positive number of employees while a large portion of the locations on the BR has an "undetermined" number of employees. The technical report by Chen and Shen (2016) proposes a calibration of the 2015 RCPM to the BR where the BR counts for "undetermined" are combined with those for businesses with less than five employees (stratum A).

6. Conclusion

For the 2015 RCPM, Bank of Canada's staff developed a flexible survey design, taking advantage of auxiliary information in the D&B database and paradata obtained during the fieldwork. As a result of the revised survey design, the SL sample fulfilled all cell targets

and cost of payment estimates from the 2015 RCPM will be more robust than the estimates in Arango and Taylor (2008b). Since the 2015 RCPM was a voluntary survey, low response rates were identified as the greatest challenge. A nonresponse analysis is recommended to address biases arising from nonresponse and frame imperfections, including follow-up with the provider of the D&B database from which the frame was build. More effort could be made to convert sample units from the initial draw into respondents or to verify their status through screening instead of sampling additional units to reach the desired response counts. In the case of large chains, their economic and statistical significance justifies additional survey effort. Bank of Canada staff found that direct outreach and networking, while labour-intensive, was more effective at obtaining their responses then using the information in the D&B data base.

Since they had access to the sampling frame, Bank of Canada staff also improved the weighting and nonresponse adjustment in the 2015 RCPM compared to the 2006 Retailer Study. The weighting, Chen and Shen (2016), and nonresponse reports, Hatko (2016), are available under a separate cover.

Acknowledgements

This work was made possible by the Bank of Canada Cost Survey Team and the Survey Technical Team: Many thanks to Heng Chen, Stan Hatko, Kim P. Huynh, Valéry D. Jiongo, Anneke Kosse, and Kyle Vincent for their continuing support and valuable comments. Casey Jones and Rallye Shen provided excellent technical support in the implementation of the clustering described in section 3.1. Casey Jones also made significant contributions to the documentation of this process. Zixin Nie spent a summer term in 2015 as intern with the Bank of Canada and dedicated his work-term report to a probabilistic model for returned survey packages. Many thanks to Currency's regional teams for their recruitment of respondents. Jean-François Beaumont (Statistics Canada), David Haziza (Université de Montreal and Statistics Canada) and Alan Roshwalb (Ipsos Reid) provided expert comments and suggestions. The author would also like to thank organizers and participants of ICES V and in particular Björn Segendorf for discussing the paper and Geoffrey Gerdes for reviewing it. The opinions rendered in this report are the author's, as are the errors and omissions.

References

ARANGO, C., AND V. TAYLOR (2008a): "Merchant Acceptance, Costs, and Perceptions of Retail Payments: A Canadian Survey," Discussion Papers 08-12, Bank of Canada.

——— (2008b): "Merchants' Costs of Accepting Means of Payment: Is Cash the Least Costly?," *Bank of Canada Review*, 2008(Winter), 15–23.

BEAUMONT, J.-F., C. BOCCI, AND D. HAZIZA (2014): "An Adaptive Data Collection Procedure for Call Prioritization," *Journal of Official Statistics*, 30(4), 607–621.

CHEN, H., C. HENRY, K. P. HUYNH, R. Q. SHEN, AND K. VINCENT (2016): "Measuring consumer cash holdings: Lessons from the 2013 Bank of Canada Methods-of-Payment survey," *Survey Practice*, 9(3).

CHEN, H., AND R. Q. SHEN (2016): "2015 Cost of Payments Study: Calibration for Single location Retailers," Bank of Canada Staff Technical Report, forthcoming.

DE MUNNIK, D., D. DUPUIS, AND M. ILLING (2013): "Assessing the Accuracy of Non-random Business Conditions Surveys: A Novel Approach," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176, 371–388.

DEMING, E., AND F. STEPHAN (1940): "On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known," *Annals of Mathematical Statistics*, 11(4), 427–444.

DILLMAN, D. A. (2006): "Why Choice of Survey Mode Makes a Difference," *Public Health Reports*, 1(121), 11–13.

DILLMAN, D. A., J. D. SMYTH, AND L. M. CHRISTIAN (2008): *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. Wiley Publishing, 3 edn.

ECDGC (2015): "Survey on merchants' cost of processing cash and card payments," Discussion paper, European Commission Directorate-General for Competition.

HATKO, S. (2016): "Nonresponse in 2015 Retailer Survey on the Cost of Payment Methods," Bank of Canada Staff Technical Report, forthcoming.

HENRY, C., K. HUYNH, AND R. SHEN (2015): "2013 Methods-of-Payment Survey Results," Discussion paper.

JIONGO, V. D. (2016): "Retailer Cost of Payment Survey 2015 Data Exploration for the Headquarters and Single Location C," Bank of Canada Staff Technical Report, forthcoming.

LEVY, P. S., AND S. LEMESHOW (2008): *Sampling of Populations: Methods and Applications*, Wiley Series in Survey Methodology. Wiley.

LINEBACK, J. F., AND K. J. THOMPSON (2010): "Conducting Nonresponse Bias Analysis for Business Surveys," in *Proceedings of the Survey Research Methods Section, ASA*.

LOHR, S. L. (2011): "Alternative survey sample designs: Sampling with multiple overlapping frames," *Survey Methodology*, 37(2).

RIVIÈRE, P. (2002): "What Makes Business Statistics Special?," *International Statistical Review*, 70(1), 145–159.

SCHMIEDEL, H., G. KOSTOVA, AND W. RUTTENBERG (2013): "The social and private costs of retail payment instruments: a European perspective," *The Journal of Financial Market Infrastructures*, 2(1), 37–75.

STATISTICS CANADA (2010): "A Brief Guide to the Business Register," `http://www23.statcan.gc.ca/imdb-bmdi/document/1105_D2_T1_V3-eng.htm`, Online; accessed 18-September-2015.

——— (2014): "Retail Trade Survey (Monthly) (MRTS)," `http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=2406`, Online; accessed 21-September-2014.

STEWART, C., I. CHAN, C. OSSOLINSKI, D. HALPERIN, AND P. RYAN (2014): "The Evolution of Payment Costs in Australia," RBA Research Discussion Papers, Reserve Bank of Australia.

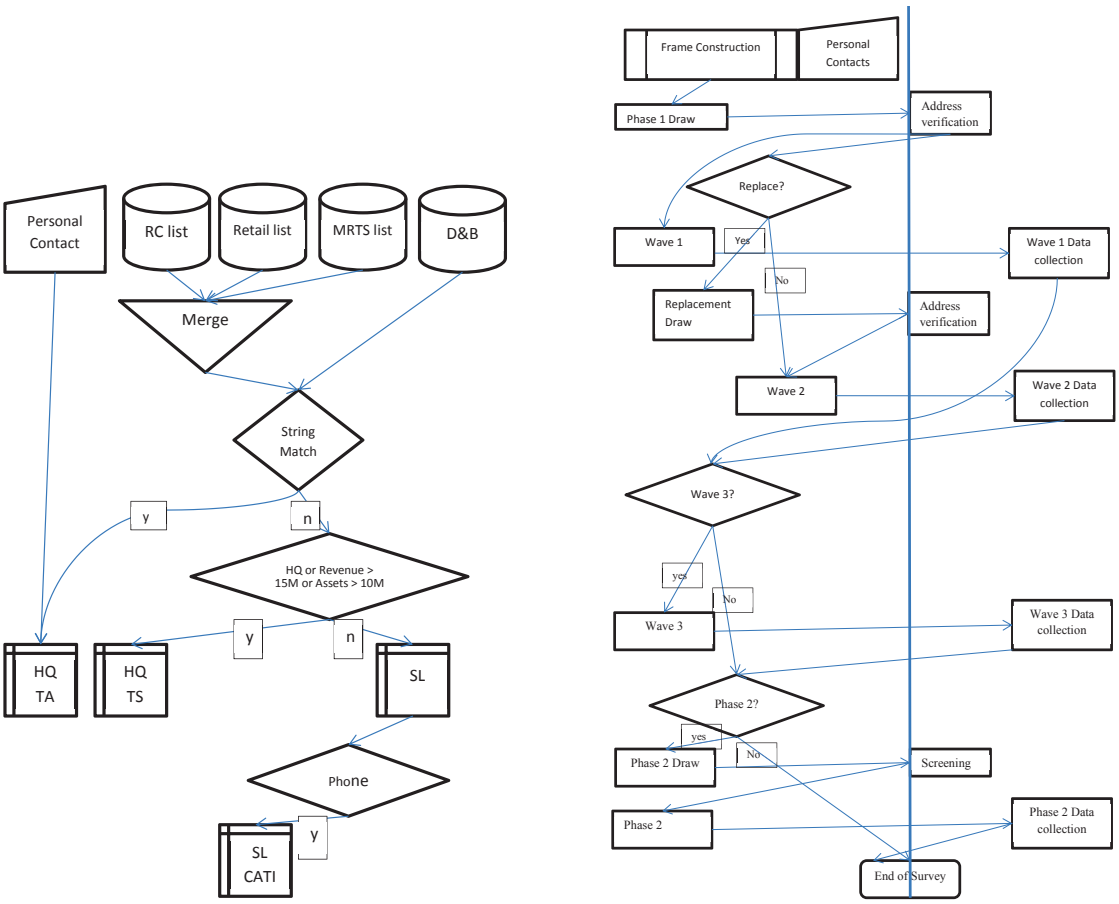

Figure 1: Sampling Frame Construction and SL Sampling

Stratum	mean	SD	min	p5	p50	p95	max
A	28.05	44.64	1.00	3.70	21.45	51.45	1114.06
B	44.84	101.29	1.40	3.28	16.27	314.67	515.75
AT	9.92	17.90	1.00	1.02	4.09	43.26	129.77
BC	24.00	54.18	2.36	2.67	15.47	106.38	332.64
ON	47.38	83.10	5.01	5.69	24.60	278.51	1114.06
PR	23.29	46.08	3.16	3.28	15.65	72.10	1037.73
QC	30.09	69.35	3.33	4.54	21.45	46.24	515.75
44	33.36	54.00	3.38	7.19	23.99	43.26	1114.06
45	14.79	16.30	1.40	3.54	9.08	22.54	220.37
72	53.95	122.68	1.00	2.67	7.38	463.46	515.75
81	26.44	33.73	3.75	4.15	21.45	51.45	1037.73

Table 3: Distribution of design weights by stratum: Columns show the mean, standard deviation, 5, 50, 95 percentiles, minimum and maximum design weight for the population

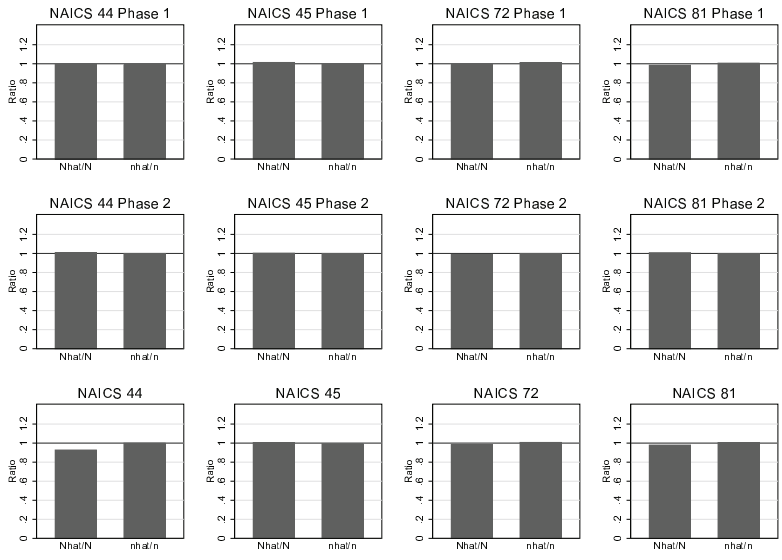

Figure 2: Sample size and population size identities-fit by NAICS stratum