

# Survey on Financing of Small And Medium Enterprises

# Methodology

**Prepared by:**
**Business Survey Methods Division**

**Survey on Financing of Small and Medium Enterprises, 2000**
**Methodology Report**

## 1 - BACKGROUND INFORMATION

As a result of recommendations made by the MacKay Task Force on the Future of the Canadian Financial Services Sector, Finance Canada, Industry Canada and Statistics Canada have been mandated to initiate a new statistical program to collect information on small and medium business financing. Statistics Canada is surveying these enterprises from across the country to determine the kind of debt and equity financing Small and medium sized businesses are relying on as well as to gain information on attempts to obtain new financing. The Survey on Financing of Small and Medium Enterprises is often referred to as the "demand-side" survey on financing of small and medium enterprises (SMEs). An independent survey on the financing of businesses from the supply-side point of view is also being conducted by Statistics Canada on a yearly basis. This paper only discusses the demand-side survey.

In 2000, a pilot survey of some 2,000 businesses was conducted for the reference year 1999 to assess the feasibility of the survey approach to collecting such information as well as produce a few national estimates for some of the key questions of interest. As a result of this pilot, many recommendations were made for the implementation of a larger scale production survey. First, significant changes were made to the questionnaire to make it simpler and shorter. Second, some modifications were made as to how the population of SMEs should be defined. Finally, the results of the pilot were used in designing the sample for the production survey. This document outlines the methodology used for conducting the larger scale production survey on financing of SMEs in reference year 2000..

**Note to Users**
The survey on financing of small and medium entreprises of 2000 was conducted for firms that were active during the survey period. The survey results did not take into account firms that may have ceased oprerations due to the lack of financing just prior to survey taking.The entrepreneurs that may have tried to launch their business and failed to do so during this survey process were also not targeted by this survey.

## 2 – SAMPLE DESIGN

### 2.1 Sampling Frame

The starting point for defining the target population for the survey of SMEs was to include all enterprises who were on the Business Register (BR) Universe file dated December 31$^{st}$, 2001. This database is constructed and adopted, using various types of tax

records from the Canada Customs and Revenue Agency and is also updated regularly, using feedback from other business surveys. The BR contains the universe of enterprises in Canada. Because of the time lag before new businesses make it onto the BR and because of the fact that young businesses are of particular interest for this survey, the frame was later supplemented by adding enterprises that were born after the creation of the initial frame. Once the universe file was created, some enterprises were removed based on auxiliary information that was available from the BR. The following enterprises were excluded from the population:

1- Enterprises with 500 or more employees
2- Enterprises with over $50 million in gross revenue
3- Enterprises coded as being non-profit (schools, hospitals, charities, ..)
4- Cooperatives
5- Joint ventures
6- Municipal/Federal Government
7- Other industries (NAICS) for which financing is not an of interest (see table 1 for list of all NAICS excluded).

Note that other exclusions such as subsidiaries could not be identified based on the frame information. Such companies were screened out at the collection phase. The final sampling frame contained 1,706,651 enterprises.

**Table 1: List of NAICS 2 and NAICS 4 excluded from SME population**

| NAICS | Description(s) |
|---|---|
| 22 | Utilities |
| 52 | Finance and Insurance |
| 55 | Management of Companies and Enterprises |
| 61 | Educational Services |
| 91 | Public Administration |
| 5321 | Need description |
| 5324 | Need description |
| 6214 | Out-Patient Care Centres |
| 6215 | Medical and Diagnostic Laboratories |
| 6219 | Other Ambulatory Health Care Services |
| 6221 | General Medical and Surgical Hospitals |
| 6222 | Psychiatric and Substance Abuse Hospitals |

| | |
|---|---|
| 6223 | Specialty (except Psychiatric and Substance Abuse) Hospitals |
| 6242 | Community Food and Housing, and Emergency and Other Relief Services |

## 2.2 Stratification of the Frame

The list frame was stratified according to Finance and Industry Canada's needs and to methodology's recommendations. Since estimates were required by region, industry type, size and age of business, these four variables were used to develop an initial stratification. Number of employees in the enterprise was used to define the size of a business and the age of the business was estimated using the date that the business was birthed on the Business Register. This is only a proxy for the actual age of the business but it was the most accurate information available. Note that the number of employees and the age of business are information collected during the interview and a reclassification based on survey results is done for tabulation purposes. The categories used for each of the stratification variables and the population counts for each category are given in the tables below

**Table 2a: Population counts by region**

| Region | Population count |
|---|---|
| Atlantic | 106,578 |
| Quebec | 403,540 |
| Ontario | 589,116 |
| Manitoba and Saskatchewan | 137,047 |
| Alberta and Northwest Territories | 222,460 |
| British Columbia and Yukon | 247,910 |

**Table 2b: Population counts by number of employees**

| Number of employees | Population count |
|---|---|
| 0 | 828,912 |
| 1-4 | 543,443 |
| 5-19 | 239,706 |
| 20-99 | 85,679 |
| 100-499 | 8,911 |

**Table 2c: Population counts by industry**

| Industry | Population count |
|---|---|
| Agriculture | 163,224 |
| Primary | 39,421 |
| Manufacturing | 89,815 |
| Wholesalers/Retailers | 187,367 |
| Professional | 177,285 |
| Knowledge Based | 78,433 |
| All other | 971,106 |

**Table 2d: Population counts by age of business**

| Age of business | Population count |
|---|---|
| Less 1 year old | 163,281 |
| 1 year old or older | 1,543,370 |

Knowledge-based industries (KBIs) include firms in a number of technology sectors, such as telecommunications carriers, video production, and computer services. It is a re-grouping of Statistics Canada's standard industry categories that is often used by Industry Canada and other organizations. Table 3 lists NAICS codes used for identifying KBIs. "Other industries" is more of a catch-all category that includes everything not included above. Although estimates for that category are not a primary goal, it must be well represented in the sample in order to have good overall estimates for all combined industries.

Finally, each of the cells defined above were further stratified in two revenue strata (low and high revenue). The purpose of this is to optimize the sample design in order to produce quantitative estimates from the collected data. Statistics Canada's Generalized Sampling System (GSAM) was used to define the low and high revenue cut-off for all of those strata. The cumulative root f rule was used for that purpose (see Sampling Techniques by Cochran, pages 128-131, 1977). Note that the revenue cutoff within each stratum will not necessarily be the same. Also, although the high revenue strata will tend to have a higher sampling fraction, they will not necessarily be take-all strata. Once all stratification was completed, the sampling frame had a total of 807 strata.

**Table 3: KBI Defined by Industry Canada**

| NAICS | DESCRIPTION |
|---|---|
| 325410 | Pharmaceutical and Medical Manufacturing |
| 333310 | Commercial and Service Industry Machinery Manufacturing |
| 334110 | Computer and Peripheral Equipment Manufacturing |
| 334210 | Telephone Apparatus Manufacturing |
| 334220 | Radio and Television Broadcasting and Wireless Communications Equipment Manufacturing |
| 334290 | Other Communications Equipment Manufacturing |
| 334310 | Audio and Video Equipment Manufacturing |
| 334410 | Semiconductor and other Electronic Component Manufacturing |
| 334511 | Navigational and Guidance Instruments Manufacturing |
| 334512 | Measuring, Medical and Controlling Devices Manufacturing |
| 335920 | Communication and Energy Wire and cable manufacturing |
| 336410 | Aerospace Product and Parts Manufacturing |
| 511210 | Software Publishers |
| 512110 | Motion Picture and Video Production |
| 512190 | Post-Production and Other Motion Picture and Video Industries |

| | |
|---|---|
| 513210 | Pay and Specialty Television |
| 513220 | Cable and Other Program Distribution |
| 513310 | Wired Telecommunication Carriers |
| 513320 | Wired Telecommunication Carriers (except satellite) |
| 513330 | Telecommunications Resellers |
| 513340 | Satellite Telecommunications |
| 513390 | Other Telecommunications |
| 514210 | Data Processing Services |
| 541360 | Geophysical Surveying and Mapping Services |
| 541370 | Surveying and Mapping (except Geophysical) Services |
| 541510 | Computer Systems Design and Related Services |
| 541620 | Environmental Consulting Services |
| 541690 | Other Scientific and technical Consulting Services |
| 541710 | Research and Development in the Physical, Engineering and Life Sciences |
| 541990 | All other Professional, Scientific and Technical Services |

### 2.3 Allocation of the sample

Based on overall budget available for the survey and studies done as a result of the pilot, the national sample size for the survey of SMEs was determined to be in the neighborhood of 25,000 enterprises. Given that estimates were required for both qualitative and quantitative type of questions, the allocation had to be flexible enough to satisfy both of these requirements. Again, GSAM was used to obtain an initial allocation, with some adjustments made subsequently to adjust for non-response, out-of-scopes and stratum jumpers. Given the sample size constraint and as a result of consultation with the clients, the following allocation strategy was used. In the text that follows, the term "domain" refers to the level at which we wish to publish estimates. For qualitative type estimates (percentages), a standard error of 2.5% was targeted for all domains consisting of the number of employees cross-tabulated with the other three main stratification variables (region, age and industry). To be conservative, the allocation method assumed a percentage equal to 50% within all the domains when determining the sample size (the standard error for a percentage is maximized when the estimate is at 50%). There was one exception for the target standard error for strata "all other industries". The constraint of 2.5% was relaxed (to 5%) because this category is not of major importance to the client. However, as previously noted, it still needed to be well represented in order to get good overall estimates for all industries. For the quantitative variables, a coefficient of variation of 5% was targeted (using the revenue variable from the frame) for the main stratification categories only. The program minimizes the overall sample size given the CV requirements described above for the various domains. No cross-tabulations were used in this allocation for quantitative data. It is therefore expected that results for qualitative variables will be publishable at a more detailed level then the financial data. Since two independent allocations were performed, the initial allocated sample size within each stratum was the maximum of those two allocations. This initial allocation resulted in a total sample size needed of 17,500 enterprises, well under the target of 25,000 units.

Some adjustments were made to increase to total sample :

1- within a stratum, a minimum sample size had to be satisfied to ensure that enough units would represent the stratum, considering non-response. The limit was set at 5.

2- Within a stratum, a maximum sampling weight (population size divided by sample size) of 500 to ensure that no single unit would represent a too important portion of the population at the estimation stage. This had been a problem with the pilot survey, especially in the 0 employee strata where there are many enterprises.

3- Sample sizes for strata where age of business is “less than 1 year old” were multiplied by a factor of 2.5. This was necessary because of the importance of having publishable results for that domain and because, based on the pilot, more than half of the businesses identified as “less than 1 year old” on the BR ended up actually being “1 year or older”.

4- Sample sizes for strata with 0 employees were multiplied by a factor of 1.2. This was to account for high incidence of stratum jumpers and the fact that death rates of businesses in that domain are higher than for businesses that have employees.

5- Finally, once steps 1 to 4 had been completed, all strata sample sizes were boosted by 20% (where this was possible) to account for non-response and out-of-scopes.

The final sample size allocated for the survey was 27,202 enterprises, slightly higher than the targeted 25,000 but still deemed acceptable.

### 2.4 Sample selection

Within each stratum, simple random sampling was used to select the units. However, one constraint was to try and minimize the overlap with the Unified Enterprise Survey (UES). When selecting a random sample of size n from a stratum, we usually assigns a random number to all units in the stratum, then sort them in ascending order and finally select the first n units in the sorted data set as being in-sample. In order to minimize the overlap, we used the same random numbers that the UES had generated for enterprises that were on both frames. A random number was generated for all units that were unique to the SME survey using the same algorithm that UES developed to generate their numbers. The actual sample overlap was minimized by sorting the resulting file in descending order instead of ascending order before selecting the sample. Using this approach resulted in an overlap of less than 1000 units between the two surveys.

Once the final sample had been selected, it was further evaluated for overlap with 2 other surveys, the pilot survey on financing of SMEs and the “attitude” survey conducted by an outside firm on behalf of Industry Canada that deals with financing of SMEs as well. The overlap evaluation with the latter survey was not exact since it was not a statistics Canada

survey and therefore did not have the same unique identifiers as our survey. It was found that very few enterprises overlapped with the pilot and an estimated 167 units overlapped with the "attitude" survey. These 167 units were assigned to experienced interviewers who were aware the companies were currently surveyed as part of a similar study. Of the 27,202 enterprises selected, 810 were identified by Business Register Division (BRD) as being out-of-business even before collection started. Therefore, only 26, 382 units were sent out on the field.

## 3 – COLLECTION RESULTS

Collection for this survey was done in two parts. In part 1, most qualitative type questions concerning the businesses latest financing requests were collected using a CATI instrument. For part 2, a questionnaire that collects detailed financial information on liabilities and balance sheet was mailed or/and faxed-out to all businesses that responded to part 1.. Telephone follow-up was used to increase response rates.

The reference period for the Part 1 of the survey coincided with the calendar year ending December 31st, 2000. The reference period for Part 2 coincided with the Fiscal year ending no later than March 31st 2001. Part 1 CATI Interiews started mid-June 2001 and ended in October 2001. Questionnaire collection for Part 2 started by end of August and ended February 2002.

A summary of collection results for part 1 is given in tables 4 and 5 that follow.

**Table 4: Response rate for Part 1 by size**

| | In scope rate (%) | Response rate (%) |
|---|---|---|
| Overall | 68 | 62 |
| 0 employees | 59 | 60 |
| 1-4 employees | 75 | 64 |
| 5-19 employees | 76 | 64 |
| 20-99 employees | 69 | 63 |
| 100-499 employees | 54 | 55 |

**Table 5: Collection results for part 1**

| |
|---|
| Sample sent on the field: 26,382 enterprises |
| Overlap with Attitude survey : 173 |
| 8,440 enterprises out of scope |

| | |
|---|---|
| -Unable to locate: | 4,463 |
| -Out of business: | 737 |
| -Screened out: | 3,240 |
| We were left with: 17,769 in scope enterprises | |
| -Refusal: | 4,003 |
| -Unable to contact: | 2,751 |
| -Completed interviews: | 11,015 |

## 4 – DATA PROCESSING

Processing of the data for the two parts of this survey was done separately so that part 1 estimates could be released before the end of 2001. All data editing and cleaning was done by SBSS Division. For part 1, no imputation was performed for missing data. Because part 1 deals strictly with qualitative data, the incidence of missing data was very small. All estimates from part 1 released were percentages of qualitative categories. For records from Part 2 of the survey, a nearest neighbor imputation system was implemented. This method of imputation involves locating a donor(from the entire «deck») of similar size and characteristics to impute data for missing or incomplete information

## 5 – ESTIMATES

For producing estimates, Statistics Canada's Generalized Estimation System (GES) was used. Initial sample weights were adjusted to account for refusals and other non-response. Estimates were produced for over 100 domains of interest defined based on stratification variables (e.g. region, industry) as well questionnaire variables (e.g. number of employees, type of creditor). Furthermore, a post-stratified estimator was used to calibrate to a known total number of enterprises in each of the number of employee categories. These known counts were obtained from the BR, taking into account death rates and out-of-scope rates.

Finally, for every estimate produces by GES, a standard error which gives the sampling error of this estimate was also obtained. An estimate was judge to be of good quality if it's standard error was small enough. The following rules based on the standard error were used to assign a measure of quality to all of the estimates of percentages.

| Standard error | Quality code |
|---|---|
| 0 - 0.025 | A – Excellent |
| 0.025-0.05 | B – Good |
| 0.05 – 0.075 | C – average |
| 0.075 – 0.10 | D – mediocre |
| 0.10 – 0.125 | E – poor, use with caution |
| > 0.125 | F – poor, use with caution |