Survey on Financing of Small And Medium Enterprises 2001

Methodology

Prepared by:
Business Survey Methods Division

Survey on Financing of Small and Medium Enterprises, 2001
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 target population of SMEs should be defined. Finally, the results of the pilot were used in designing the sample for the production surveys for 2000 and 2001.

This document outlines the methodology used for conducting the production survey on financing of SMEs in reference year 2001.

Note to Users
The survey on financing of small and medium enterprises of 2001 was conducted for firms that were active during the survey period. The survey results did not take into account firms that may have ceased operations 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 31st, 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-   Affiliates more than 50% owned by another society
7-   Municipal/Federal Government
8-   Other industries (NAICS) for which financing is not 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 1,833,821 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	Automotive Equipment Rental and Leasing
5324	Machinery and Equipment Rental and Leasing
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	109 514
Quebec	417 855
Ontario	647 407
Manitoba and Saskatchewan	144 227
Alberta and Northwest Territories	246 440
British Columbia and Yukon	268 378

Table 2b: Population counts by number of employees

Number of employees	Population count
0	994 646
1-4	518 879
5-19	225 432
20-99	83 258
100-499	11 606

Table 2c: Population counts by industry

Industry	Population count
Agriculture	171 080
Primary	42 148
Manufacturing	93 082
Wholesalers/Retailers	308 924
Professional	199 723
Knowledge Based	95 397
All other	923 467

Table 2d: Population counts by age of business

Age of business	Population count
Less 1 year old	263 114
1 year old or older	1 570 707

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 8,000 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 main stratification variables region, age, number of full-time employees 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%).

Unlike what was done for the 2000 survey, the coefficient of variation of quantitative variables were not targeted in the allocation of the sample.

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, so that the stratum size must be greater than 5 units.

2- Within a stratum, a maximum sampling weight (population size divided by sample size) of 1,500 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.

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. The sample sizes for strata with 100 to 499 employees were multiplied by a factor of 1.5

to account for enterprises for which the number of employees increased and being excluded from the target population of enterprises with less then 500 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 8,523 enterprises, slightly higher than the targeted 8,000.

2.4 Sample selection

Within each stratum, simple random sampling was used to select the units. No overlap studies were undertaken for the 2001 survey.

Of the 8,523 enterprises selected, 284 were identified by Business Register Division (BRD) as being out-of-business even before collection started. Therefore, only 8,239 units were sent out on the field.

3 – COLLECTION RESULTS

This survey consisted mostly in qualitative type questions concerning the businesses latest financing requests. The data were collected using a CATI instrument corresponding to Part 1 of the 2000 survey. The reference period for the Survey of Financing of Small and Medium-sized Enterprises for 2001 coincided with the calendar year ending December 31st, 2001. Telephone interviews started in September and ended October 2002.

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

Table 4a: Response rate by size

Category of number of employees	Response rate (%)
Overall	**66.3**
0 employees	62.5
1-4 employees	68.7
5-19 employees	70.5
20-99 employees	68.2
100-499 employees	65.3

Table 4b: Response rate by region

Region	Response rate (%)
Overall	**66.3**
Atlantic	71.7
Quebec	77.5
Ontario	64.6
Manitoba and Saskatchewan	59.1
Alberta and Northwest Territories	65.8
British Columbia and Yukon	59.2

Table 4c: Response rate by industry

Industry	Response rate (%)
Agriculture	59.1
Primary	67.8
Manufacturing	72.3
Wholesalers/Retailers	67.9
Professional	68.9
Knowledge Based	71.9
All other	62.1

Table 4d: Response rate by age of business

Age of business	Response rate (%)
Less 1 year old	68.8
1 year old or older	64.5

Table 5: Collection results

Sample sent on the field: 8,523 enterprises
 	
2,715 enterprises out of scope
     -Unable to locate: 1,200
     -Out of business: 687
     -Screened out: 828
We were left with: 5,808 in scope enterprises
     -Refusal: 760
     -Unable to contact: 1,206
     -Completed interviews: 3,842

4 – DATA PROCESSING

The complete records were verified by the Small Business and Special Survey division (SBSS) to detect inconsistencies and outliers. From Business Survey Methods Division (BSMD), a system for edits verifying invalid values and verifying skips as well as substitution of missing values based on logical imputation rules (ex: pro-rating of totals based on information obtained) was implemented.

Furthermore, a nearest neighbor imputation system was used to correct partial non-response present in the returned questionnaires. This imputation method consists in replacing one or more missing values from a respondent, called receiver, by values provided by one or more respondents, called donors. Nearest donor(s) are determined based on size (revenue and number of employees closest to the receiver's) and other characteristics (industry and type of financing requested identical to those of the receiver). The values selected for imputation must pass pre-established edit rules (post-imputation edits).

Table 6: Average Imputation rates by section (%)

Global *	2.22 *
Section B	1.76
Section C	1.10
Section D	0.89
Section E	1.09
Section F	0.54
Section G *	2.04 *
Section G (including question G.3)	(27.38)
Section H	8.09

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, number of employees, age of business) as well questionnaire variables. 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 the goal of this calibration is to obtain population estimates (number of enterprises) corresponding to BR counts.

Finally, for every estimate produces by GES, a quality measure easily interpreted was computed. For totals, this measure consists in the coefficient of variation (C.V.) while for proportions, the measure used is the standard error (which gives the sampling error of this estimate). An estimate was judged to be of good quality if it's C.V. or standard error and

imputation rate were 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

NB The measures of quality do not account for the imputation rate. Since imputation rates are usually small, the standard error gives a good image of the estimate's quality.

Confidentiality

To protect respondents' confidentiality, estimates obtained from four observations or less in a domain of interest were not published.