

SURVEY OF SELF-EMPLOYMENT

USER GUIDE

January 2002



Statistics
Canada

Statistique
Canada

Canada

Table of Contents

1.0	Introduction	5
2.0	Background	7
3.0	Objectives	9
4.0	Concepts and Definitions	11
	4.1 Labour Force Survey Concepts and Definitions	11
	4.2 Survey of Self-employment Concepts and Definitions	13
5.0	Survey Methodology	15
	5.1 Population Coverage	15
	5.2 Sample Design	15
	5.2.1 Primary Stratification	15
	5.2.2 Types of Areas	16
	5.2.3 Secondary Stratification	16
	5.2.4 Cluster Delineation and Selection	17
	5.2.5 Dwelling Selection	17
	5.2.6 Person Selection	17
	5.3 Sample Size	19
	5.4 Sample Rotation	19
	5.5 Modifications to the LFS Design for the Supplement	19
	5.6 Sample Size by Province for the Supplement	19
6.0	Data Collection	21
	6.1 Interviewing for the LFS	21
	6.2 Supervision and Control	22
	6.3 Non-Response to the LFS	22
	6.4 Data Collection Modifications for the Survey of Self-employment	22
	6.5 Non-Response to the Survey of Self-employment	22
7.0	Data Processing	23
	7.1 Data Capture	23
	7.2 Editing	23
	7.3 Coding of Open-ended Questions	23
	7.4 Creation of Derived Variables	23
	7.5 Weighting	24
	7.6 Suppression of Confidential Information	24
8.0	Data Quality	27
	8.1 Response Rates	27
	8.2 Survey Errors	28
	8.2.1 The Frame	28
	8.2.2 Data Collection	29
	8.2.3 Data Processing	29
	8.2.4 Non-response	29
9.0	Guidelines for Tabulation, Analysis and Release	33
	9.1 Rounding Guidelines	33
	9.2 Sample Weighting Guidelines for Tabulation	33
	9.2.1 Definitions of types of estimates: Categorical vs. Quantitative	34

2000 Survey of Self-employment – User Guide

	9.2.2	Tabulation of Categorical Estimates	35
	9.2.3	Tabulation of Quantitative Estimates	35
9.3		Guidelines for Statistical Analysis	35
9.4		C.V. Release Guidelines	36
10.0		Approximate Sampling Variability Tables	39
	10.1	How to use the C.V. tables for Categorical Estimates	39
		10.1.1 Examples of using the C.V. tables for Categorical Estimates	41
	10.2	How to use the C.V. tables to obtain Confidence Limits	43
		10.2.1 Example of using the C.V. tables to obtain confidence limits	44
	10.3	How to use the C.V. tables to do a t-test	45
		10.3.1 Example of using the C.V. tables to do a t-test	45
	10.4	Coefficients of Variation for Quantitative Estimates	45
	10.5	Release cut-off's for the Survey of Self-employment	46
	10.6	C.V. Tables	47
11.0		Weighting	57
	11.1	Weighting Procedures for the LFS	57
	11.2	Weighting Procedures for the Survey of Self-employment	58
12.0		Questionnaires	61
	12.1	The Labour Force Survey Questionnaire	61
	12.2	The Survey of Self-Employment Questionnaires	61
13.0		File Layout with Univariate Frequencies	63

1.0 Introduction

The Survey of Self-employment (SSE) was conducted by Statistics Canada in April 2000 with the cooperation and support of Human Resources Development Canada (HRDC). This manual has been produced to facilitate the manipulation of the microdata file of the survey results.

Any questions about the data set or its use should be directed to:

Statistics Canada

Client Services
Special Surveys Division
Tel: (613) 951-7355 OR 1-888-297-7355
Fax: (613) 951-3012
ssd@statcan.ca

Elizabeth Majewski
Special Surveys Division,
D6 Section
5th floor, Jean Talon Building
Tunney's Pasture
Ottawa, Ontario K1A 0T6
(613) 951-4584
[e-mail: Elizabeth.Majewski@statcan.ca](mailto:Elizabeth.Majewski@statcan.ca)

Human Resources Development Canada
Benoit Delage
Applied Research Branch
Promenade du Portage
Phase II, 7th Floor
Hull, Quebec K1A 0J9
(819) 953-8119
e-mail: benoit.delage@hrdc-drhc.gc.ca

2.0 Background

Strong growth in the number of self-employed Canadians that was observed in the 1990s stimulated interest in self-employment. Among the existing sources of information on this topic, the most extensive is Statistics Canada publication *The Self-employed* (Autumn 1997) which portrayed the self-employed using data from the LFS, the annual Survey of Consumer Finances, and the 1995 Survey of Work Arrangements. It provided a thorough picture of the basic socio-demographic characteristics of this population, but due to the lack of data it did not cover several specific aspects of self-employment. Human Resources Development undertook to enrich the data sources on self-employment by funding a survey devoted entirely to this topic.

3.0 Objectives

The primary objective of the survey is to provide a profile of those who are self-employed in their main job. Specific issues covered include:

- motivation behind becoming self-employed, i.e. “recession or structural push” (self-employment as a result of poor chances of finding suitable paid-employment) versus “entrepreneurial pull ” (self-employment as a result of preference for “being own boss”);
- differences in the socio-economic characteristics of those who were pushed and those who were pulled;
- reasons for growth of own account self-employment, in particular changing employment relationships;
- extent and sources of benefit coverage;
- incidence of formal training;
- barriers to training;
- informal training as an alternative to formal training;
- respondents’ perceptions of financial stability;
- mechanisms for dealing with personal financial difficulties;
- interest in having an insurance program providing benefits in the event poor business conditions cause personal financial problems; and
- membership in professional, occupational or trade associations.

4.0 Concepts and Definitions

This chapter outlines concepts and definitions of interest to the users. The concepts and definitions used in the Labour Force Survey are described in section 4.1 while those specific to the Survey of Self-employment are given in section 4.2. Users are referred to Chapter 12 of this document for a copy of the actual survey forms used.

4.1 Labour Force Survey Concepts and Definitions

Labour Force Status

Status of the respondent in the labour market : a member of the non-institutional population 15 years and over is designated as either **employed, unemployed or not in the labour force**.

Employed

Employed persons are those who, during the reference week:

- (a) did any work¹ at all
- (b) had a job but were not at work due to:
 - own illness or disability
 - personal or family responsibilities
 - bad weather
 - labour dispute
 - vacation
 - other reason not specified above (excluding persons on layoff and persons whose job attachment was to a job starting at a definite date in the future).

Unemployed

Unemployed persons are those who, during the reference week:

¹ Work includes any work for pay or profit, that is, paid work in the context of an employer-employee relationship, or self-employment. It also includes unpaid family work where unpaid family work is defined as unpaid work which contributed directly to the operation of a farm, business or professional practice owned or operated by a related member of the household. Such activities may include keeping books, selling products, waiting on tables, and so on. Tasks such as housework or maintenance of the home are not considered unpaid family work.

- (a) were without work, had actively looked for work in the past four weeks (ending with reference week), and were available for work²;
- (b) had not actively looked for work in the past four weeks but had been on layoff³ and were available for work;
- (c) had not actively looked for work in the past four weeks but had a new job to start in four weeks or less from the reference week, and were available for work.

Not in the Labour Force

Those persons in the civilian non-institutional population 15 years of age and over who, during the reference week, were neither employed nor unemployed.

Industry and Occupation

The Labour Force Survey provides information about the occupation and industry attachment of employed and unemployed persons, and of persons not in the labour force who have held a job in the past five years. Since 1984, these statistics have been based on the 1980 Standard Occupational Classification and the 1980 Standard Industrial Classification. Prior to 1984, the 1971 Standard Occupational Classification and the 1970 Standard Industrial Classification were used.

Reference week

Entire calendar week covered by the Labour Force Survey each month. It is usually the week containing the 15th day of the month. The interviews are conducted during the following week, called the Survey Week, and the labour force status determined is that of the reference week.

Full-time

Full-time employment consists of persons who usually work 30 hours or more per week, plus those who usually work less than 30 hours but consider themselves to be employed full-time (e.g. airline pilots).

Part-time

Part-time employment consists of all other persons who usually work less than 30 hours per week.

² Persons in this group meeting the following criteria are regarded as available:
(i) were full-time students seeking part-time work who also met condition (ii) below. (Full-time students looking for full-time work are classified as not available for work in the reference week.)
(ii) reported that there was no reason why they could not take a job in reference week, or if they could not take a job it was because of "own illness or disability", "personal or family responsibilities", or "already had a job".

³ Persons are classified as being on layoff only when they expect to return to the job from which they were laid off.

4.2 Survey of Self-employment Concepts and Definitions

The definitions of self-employment status and of start-date were adopted from the Labour Force Survey:

Self-employed - there are two types of self-employed included in the SSE:

- Working owners of incorporated businesses, farms or professional practices;
- Working owners of unincorporated businesses, farms, professional practices, and other self-employed (including those who as, for example, baby sitters, do not have a business).

Unpaid family workers are not part of this survey's target population.

Start-date of the current self-employment:

- For self-employed with business: year/month when business was created or acquired.
- For self-employed without business: year/month when the most recent period of continuous work began.

Business partner: for a partnership to exist, the business/professional practice must be registered as a partnership.

Subcontractor: there is no employer-employee relationship between the parties. They have a pre-arranged business relationship and it is the subcontractor who covers the costs of his/her benefits and taxes.

Formal training:

- has a structured content;
- usually includes an evaluation of participants resulting in a formal recognition such as a certificate, a diploma or an evaluation report;
- usually requires payment of a fee;
- is usually given in a classroom setting, but can also be offered by correspondence, through the television or through the Internet, provided the participant is officially enrolled.

Informal training:

- includes situations where there is an intention to learn certain skills or to gain certain knowledge without enrollment in a formal training;
- may involve studying manuals or other publications, as well as electronic tutorials (not as part of formal course), observation or discussion with those from whom one can learn;
- does not involve costs, other than the costs of manuals, software or trade magazines.

Learning that occurs as a natural result of practice is not considered informal training.

Income from unincorporated business:

Net income before taxes and deductions. If a partnership, only respondent's share was to be reported.

- Farmers: includes farm program payments, Canadian Wheat Board payments, crop insurance.
Excludes: operating expenses, depreciation allowances, capital cost allowances, value of food and fuel produced and consumed on own farm.
- Non-farmers: excludes operating expenses, depreciation allowances, capital cost allowances.

Income from incorporated business or professional practice:

Gross personal income before taxes and deductions.

5.0 Survey Methodology

The Survey of Self-employment was administered in April 2000 to a sub-sample of the dwellings in the Labour Force Survey (LFS) sample, and therefore its sample design is closely tied to that of the LFS. The LFS design is briefly described in Sections 5.1 to 5.4⁴. Sections 5.5 and 5.6 describe how the SSE departed from the basic LFS design in April 2000.

5.1 Population Coverage

The LFS is a monthly household survey whose sample of individuals is representative of the civilian, non-institutionalized population 15 years of age or older in Canada's ten provinces. Specifically excluded from the survey's coverage are residents of the Yukon⁵, Nunavut and Northwest Territories, persons living on Indian Reserves, full-time members of the Canadian Armed Forces and inmates of institutions. These groups together represent an exclusion of approximately 2% of the population aged 15 or over.

5.2 Sample Design

The LFS has undergone an extensive redesign, culminating in the introduction of the new design at the end of 1994. The LFS sample is based upon a stratified, multi-stage design employing probability sampling at all stages of the design. The design principles are the same for each province. A diagram summarizing the design stages appears at the end of this section.

5.2.1 Primary Stratification

Provinces are divided into economic regions and employment insurance regions. Economic regions (ERs) are geographic areas of more or less homogeneous economic structure formed on the basis of federal provincial agreements. They are relatively stable over time. Employment insurance economic regions (EIERs) are also geographic areas, and are roughly the same size and number as ERs, but they do not share the same definitions. Labour force estimates are produced for the EIER regions for the use of Human Resources Development Canada.

The intersections of the two types of regions form the first level of stratification for the LFS. These ER/EIER intersections are treated as primary strata and further stratification is carried out within them (see section 5.2.3). Note that a third set of regions, Census Metropolitan Areas (CMAs), is also respected by stratification in the current LFS design, since each CMA is also an EIER.

⁴ A detailed description of the previous LFS design is available in the Statistics Canada publication entitled **Methodology of the Canadian Labour Force Survey, 1984-1990** (catalogue #71-526).

⁵ Since 1992, the LFS has been administered in the Yukon, using an alternative methodology that accommodates some of the operational difficulties inherent to remote locales. To improve reliability due to small sample size, estimates are available on a three month average basis only. These estimates are not included in national totals.

5.2.2 Types of Areas

The primary strata (ER/EIER intersections) are further disaggregated into 3 types of areas: rural, urban, and remote areas. Urban and rural areas are loosely based on the Census definitions of urban and rural, with some exceptions to allow for the formation of strata in some areas. Urban areas include the largest CMAs down to the smallest villages categorized by the 1991 Census as urban (1,000 people or more), while rural areas are made up of areas not designated as urban or remote.

All urban areas are further subdivided into two types: those using an apartment list frame and an area frame, as well as those using only an area frame.

Approximately 1% of the LFS population is found in remote areas of provinces which are less accessible to LFS interviewers than other areas. For administrative purposes, this portion of the population is sampled separately through the remote area frame. Some populations, not congregated in places of 25 or more people, are excluded from the sampling frame.

5.2.3 Secondary Stratification

In urban areas with sufficiently large numbers of apartment buildings, the strata are subdivided into apartment frames and area frames. The apartment list frame is a register which is based upon information supplied by CMHC and is maintained in the 18 largest cities across Canada. The purpose of this is to ensure better representation of apartment dwellers in the sample as well as to minimize the effect of growth in clusters, due to construction of new apartment buildings. In the major cities, the apartment strata are further stratified into low income strata and regular strata.

Where it is possible and/or necessary, the urban area frame is further stratified into regular strata, high income strata, and low population density strata. Most urban areas fall into the regular urban strata, which, in fact, cover the majority of Canada's population. High income strata are found in major urban areas, while low density urban strata consist of small towns that are geographically scattered.

In rural areas, the population density can vary greatly from relatively high population density areas to low population density areas, resulting in the formation of strata that reflect these variations. The different stratification strategies for rural areas were based not only on concentration of population, but also on cost-efficiency and interviewer constraints.

In each province, remote settlements are sampled proportional to the number of dwellings in the settlement, with no further stratification taking place. Dwellings are selected using systematic sampling in each of the places sampled.

5.2.4 Cluster Delineation and Selection

Households in final strata are not selected directly. Instead, each stratum is divided into clusters, and then a sample of clusters is selected within the stratum. Dwellings are then sampled from selected clusters. Different methods are used to define the clusters, depending on the type of stratum.

Within each urban stratum in the urban area frame, a number of geographically contiguous groups of dwellings, or clusters, are formed based upon 1991 Census counts. These clusters are generally a set of one or more city blocks or block faces. The selection of a sample of clusters (always 6 or a multiple of 6 clusters) from each of these secondary strata represents the first stage of sampling in most urban areas. In some other urban areas, Census Enumeration Areas (EAs) are used as clusters. In the low density urban strata, a three-stage design is followed. Under this design, two towns within a stratum are sampled, and then six or 24 clusters within each town are sampled.

For urban apartment strata, instead of defining clusters, the apartment building is the primary sampling unit. Apartment buildings are sampled from the list frame with probability proportional to the number of units in each building.

Within each of the secondary strata in rural areas, where necessary, further stratification is carried out in order to reflect the differences among a number of socio-economic characteristics within each stratum. Within each rural stratum, six EAs or two or three groups of EAs are sampled as clusters.

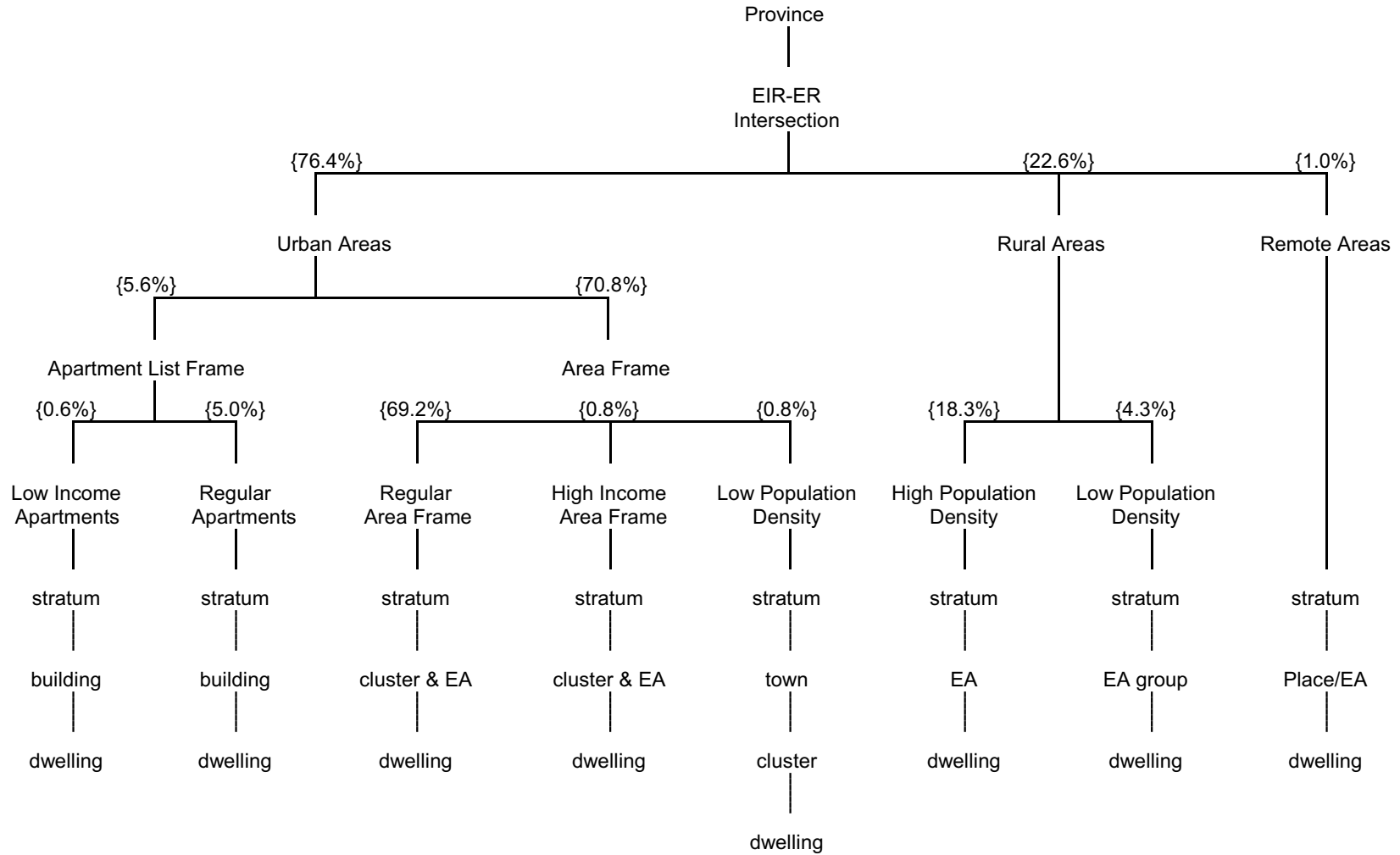
5.2.5 Dwelling Selection

In all three types of areas (urban, rural and remote areas) selected clusters are first visited by enumerators in the field and a listing of all private dwellings in the cluster is prepared. From the listing, a sample of dwellings is then selected. The sample yield depends on the type of stratum. For example, in the urban area frame, sample yields are either 6 or 8 dwellings, depending on the size of the city. In the urban apartment frame, each cluster yields 5 dwellings, while in the rural areas and EA parts of cities, each cluster yields 10 dwellings. In all clusters, dwellings are sampled systematically. This represents the final stage of sampling.

5.2.6 Person Selection

Demographic information is obtained for all persons for whom the selected dwelling is the usual place of residence. LFS information is obtained for all civilian household members 15 years of age or older. Response burden is minimized for the elderly (70 years of age or older) by carrying forward their responses for the initial interview to the subsequent five months in the survey.

Labour Force Survey Sample Design - 1995+



| = level of stratification

EIR - Employment Insurance Region

EA - Census Enumeration Area

ER - Economic Region

cluster - set of blockfaces

{%} - percentage of total sample

| = stage of sampling

5.3 Sample Size

The sample size of eligible persons in the LFS is determined so as to meet the statistical precision requirements for various labour force characteristics at the provincial and subprovincial level, to meet the requirements of federal, provincial and municipal governments as well as a host of other data users.

The monthly LFS sample consists of approximately 60,000 dwellings. After excluding dwellings found to be vacant, dwellings demolished or converted to non-residential uses, dwellings containing only ineligible persons, dwellings under construction, and seasonal dwellings, about 53,000 dwellings remain which are occupied by one or more eligible persons. From these dwellings, LFS information is obtained for approximately 102,000 civilians aged 15 or over.

5.4 Sample Rotation

The LFS employs a panel design whereby the entire monthly sample of dwellings consists of 6 panels, or rotation groups, of approximately equal size. Each of these panels is, by itself, representative of the entire LFS population. All dwellings in a rotation group remain in the LFS sample for 6 consecutive months after which time they are replaced (rotated out of the sample) by a new panel of dwellings selected from the same or similar clusters.

This rotation pattern was adopted to minimize any problems of non-response or respondent burden that would occur if households were to remain in the sample for longer than 6 months. It also has the statistical advantage of providing a common sample base for short-term month-to-month comparisons of LFS characteristics, since five of the six rotation groups in the LFS sample are common from month to month.

Because of the rotation group feature, it is possible to readily conduct supplementary surveys using the LFS design but employing less than the full size sample.

5.5 Modifications to the LFS Design for the Supplement

The Survey of Self-employment used five of the six rotation groups in the April 2000 LFS sample. For the Survey of Self-employment, the coverage of the LFS was modified to include only the eligible members of the household. To qualify for the supplement one had to be aged 15 to 69 and to be self-employed in the main job during the reference week in April 2000. Full time students and those who, on average, work fewer than 11 hours per week were excluded.

5.6 Sample Size by Province for the Supplement

The following table shows the number of household members in the LFS sampled rotations who were eligible for the SSE supplement.

PROVINCE	SAMPLE SIZE
Newfoundland and Labrador	179
Prince Edward Island	207
Nova Scotia	386
New Brunswick	263
Quebec	1,060
Ontario	1,956
Manitoba	507
Saskatchewan	676
Alberta	715
British Columbia	674
CANADA	6,623

6.0 Data Collection

Data collection for the LFS is carried out each month during the week following the LFS reference week, usually the third week of the month.

6.1 Interviewing for the LFS

Statistics Canada interviewers, who are part-time employees hired and trained specifically to carry out the LFS, contact each of the sampled dwellings to obtain the required labour force information. Each interviewer contacts approximately 70 dwellings per month.

Dwellings new to the sample are contacted through a personal visit. The interviewer first obtains socio-demographic information for each household member and then obtains labour force information for all eligible members. Provided there is a telephone in the dwelling and permission has been granted, subsequent interviews are conducted by telephone. As a result, approximately 85% of all dwellings are interviewed by telephone. In these subsequent monthly interviews, as they are called, the interviewer confirms the socio-demographic information collected in the first month and collects the labour force information for the current month.

In all dwellings, information about all household members is obtained from a knowledgeable household member - usually the person at home when the interviewer calls. Such 'proxy' reporting, which accounts for approximately 55% of the information collected, is used to avoid the high cost and extended time requirements that would be involved in repeat visits or calls necessary to obtain information directly from each respondent.

At the conclusion of the LFS monthly interviews, interviewers introduce the supplementary survey, if any, to be administered to some or all household members that month.

If, during the course of the six months that a dwelling normally remains in the sample, an entire household moves out and is replaced by a new household, information is obtained about the new household for the remainder of the six-month period.

6.2 Supervision and Control

All LFS interviewers are under the supervision of a staff of senior interviewers who are responsible for ensuring that interviewers are familiar with the concepts and procedures of the LFS and its many supplementary surveys, and also for periodically monitoring their interviewers and reviewing their completed documents. The senior interviewers are, in turn, under the supervision of the LFS program managers, located in each of the Statistics Canada regional offices.

6.3 Non-Response to the LFS

Interviewers are instructed to make all reasonable attempts to obtain LFS interviews with members of eligible households. For individuals who at first refuse to participate in the LFS, a letter is sent from the Regional Office to the dwelling address stressing the importance of the survey and the household's cooperation. This is followed by a second call (or visit) from the interviewer. For cases in which the timing of the interviewer's call (or visit) is inconvenient, an appointment is arranged to call back at a more convenient time. For cases in which there is no one home, numerous call backs are made. Under no circumstances are sampled dwellings replaced by other dwellings for reasons of non-response.

Each month, after all attempts to obtain interviews have been made, a small number of non-responding households remain. For households non-responding to the LFS and for which LFS information was obtained in the previous month, this information is brought forward and used as the current month's LFS information. No supplementary survey information is collected for these households.

6.4 Data Collection Modifications for the Survey of Self-employment

Upon completion of the Labour Force Survey interview, the interviewer asked to speak to the person(s) eligible for the Survey of Self-employment. If this person was not available, the interviewer arranged for a convenient time to phone back. Proxy response was not allowed, hence the collection period was extended by one week to allow the interviewers time to contact the individuals eligible for the supplement.

6.5 Non-Response to the Survey of Self-employment

For households responding to the LFS, the next stage of data collection was to administer the Survey of Self-employment. In total, 6,623 individuals were eligible for the supplementary survey; the SSE interview was completed for 4,023 individuals and 4,015 records qualified to be included in the survey file for a response rate of 60.62%. More detailed information on response rates is presented in Chapter 8 (Data Quality).

7.0 Data Processing

The main output of the Survey of Self-employment is a "clean" microdata file. This section presents a brief summary of the processing steps involved in producing this file.

7.1 Data Capture

The survey responses were entered during the computer assisted interviewing. The data capture program automatically followed the flow of the questionnaire and allowed for checking if the codes entered were within a valid range. Interviewers transmitted the data from their machines to the regional offices of Statistics Canada, and next to the head office.

7.2 Editing

The first type of error treated was errors in questionnaire flow, where questions which did not apply to the respondent (and should therefore not have been answered) were found to contain answers. In this case a computer edit automatically eliminated superfluous data by following the flow of the questionnaire implied by answers to previous, and in some cases, subsequent questions.

The second type of error treated involved a lack of information in questions which should have been answered. For this type of error, a non-response or "not-stated" code was assigned to the item.

7.3 Coding of Open-ended Questions

There were no open-ended questions in the SSE questionnaire. However, there were 15 partially open-ended questions with "other, specify" answers in an open-ended format. These answers were examined and either recoded or remained as "other". The recoding was done into existing or specially created answer categories.

7.4 Creation of Derived Variables

A number of data items on the microdata file have been derived by combining items on the questionnaire in order to facilitate data analysis. For each derived variable, there is a comment on the record layout stating which survey questions were used to create the variable. As an example, selected derived variables are presented below.

DVBENEF measures the benefit coverage based on questions BC_Q1, BC_Q4, and BC_Q7 and has four values: three benefits, two benefits, one benefit, and no benefits. Only positive answers are counted; the lack of response is interpreted as a lack of a benefit.

DVCHOICE distinguishes between self-employed by choice, involuntary, discouraged and adjusted self-employed. It is derived from MS_Q1 and MS_Q7. Persons who became self-employed for reason(s) other than the lack of suitable job and who would not accept a paid job (MS_Q1=2 and MS_Q7=2) are considered to be "self-employed by choice" (1), while those who became self-employed due to the absence of a suitable paid job and would accept a paid-job (MS_Q1=1 and MS_Q7=1) are classified as "involuntary self-employed" (2). Those who voluntarily became self-employed, but would now rather be paid workers (MSQ1not equal 1 and MSQ7=1) are labelled "discouraged" (3), and finally, those who originally did not choose self-employment, but would not like to leave self-employment (MSQ1=1 and MSQ7 not equal 1) are classified as "adjusted to self-employment" (4).

DVSETENU provides duration of the current self-employment in years, grouped into five ranges. It is calculated from two other derived variables: start month and start year, which were created using the LFS Q118 for non-proxy LFS respondents, and the SSE questions VL_4A- VL_Q5B together with the LFS information for the LFS proxy respondents. When the start year was provided, but the start month was missing (125 cases), January was imputed to make calculation of the duration of self-employment possible.

DVINCORP (incorporation status), DVEMPLWK (having employees in the reference week, and DVSECOWM (class of worker, which combines incorporation status with presence or absence of employees) use data from two sources. They combine the responses to the LFS questions referring to incorporation and employees obtained during non-proxy interviews with the LFS responses collected by proxy and verified by the SSE.

Start date of the current self-employment was the third variable verified with the respondents whose LFS data came from a proxy interview. The derived variables DVSTARTY, DVSTARTM and DVSTARTD were created to combine the original LFS data and the data verified by the SSE.

7.5 Weighting

The principle behind estimation in a probability sample such as the LFS is that each person in the sample "represents", besides himself or herself, several other persons not in the sample. For example, in a simple random 2% sample of the population, each person in the sample represents 50 persons in the population.

The weighting phase is a step which calculates, for each record, what this number is. This weight appears on the microdata file, and must be used to derive meaningful estimates from the survey. For example, if the number of self-employed who are required to belong to a professional, occupational or trade association is to be estimated, it is done by selecting the records referring to those individuals in the sample with that characteristic and summing the weights entered on those records.

Details of the method used to calculate these weights are presented in Chapter 11.

7.6 Suppression of Confidential Information

It should be noted that the 'Public Use' microdata files described above differ in a number of important respects from the survey 'master' files held by Statistics Canada. These differences are the result of actions taken to protect the anonymity of individual survey respondents. Users requiring access to information excluded from the microdata files may purchase custom tabulations. Estimates generated will be released to the user, subject to meeting the guidelines for analysis and release outlined in Section 9 of this document.

Region - Suppression of Geographic Identifiers

The survey master data file includes explicit geographic identifiers for province and the three largest Census Metropolitan Areas (CMAs): Toronto, Montreal and Vancouver. The survey public-use microdata file does not contain any geographic identifiers below the regional level.

LFS Variables

Several LFS variables have been further grouped on the public use file to prevent identification of respondents, for example, marital status (MARSTATG) has not six but three categories. The answers “Married” and “Living common law”, as well as “Widowed”, “Separated” and “Divorced” have been combined.

SSE Variables

The following SSE variables have been affected by the modifications made for confidentiality reasons:

VL_Q8	Spouse registered as a business partner: suppressed on 10 records.
PW_Q1	Main activity before the current self-employment : “Retired” grouped with “other”.
WA_Q1	Franchise or not: suppressed.
FN_Q12_7	Declared bankruptcy: suppressed.
FN_Q13_4	Relied on own pension to deal with financial difficulties: suppressed
FQ_Q1	Born in Canada or not: suppressed on all the records in Atlantic region.
FQ_Q2	Year of immigration: suppressed.
DVSTARTY	Start year of current self-employment: collapsed 1942-1954.
DVINCOME	Capped at \$75,000, suppressed on 225 records.

8.0 Data Quality

8.1 Response Rates

The following table summarizes the response rates to the Labour Force Survey and to the Survey of Self-employment.

	Household response rate for full LFS in April, 2000 (*1)	Household response rate for LFS rotations 1,2,3,5,6 (*1)	Number of respondents in SSE	Person response rate to SSE (*2)
Newfoundland and Labrador	93.5	94.3	107	59.8%
Prince Edward Island	92.6	93.6	106	51.2%
Nova Scotia	91.2	92.6	230	59.6%
New Brunswick	90.9	91.3	150	57.0%
Quebec	92.6	93.3	724	68.3%
Ontario	90.9	91.7	1162	59.4%
Manitoba	95.9	96.3	315	62.1%
Saskatchewan	95.8	96.1	410	60.6%
Alberta	96.4	96.3	395	55.2%
British Columbia	89.3	90.6	416	61.7%
CANADA	92.5	93.1	4015	60.6%

Note:

- (*1) Response rate is number of responding households as a percentage of number of eligible households.
- (*2) Response rate is number of individuals responding to SSE as a percentage of number of individuals eligible for SSE and responding to LFS in rotations sampled .

8.2 Survey Errors

The estimates derived from this survey are based on a sample of households. Somewhat different figures might have been obtained if a complete census had been taken using the same questionnaire, interviewers, supervisors, processing methods, etc. as those actually used. The difference between the estimates obtained from the sample and the results from a complete count taken under similar conditions is called the sampling error of the estimate.

Errors which are not related to sampling may occur at almost every phase of a survey operation. Interviewers may misunderstand instructions, respondents may make errors in answering questions, the answers may be incorrectly entered on the questionnaire and errors may be introduced in the processing and tabulation of the data. These are all examples of non-sampling errors.

Over a large number of observations, randomly occurring errors will have little effect on estimates derived from the survey. However, errors occurring systematically will contribute to biases in the survey estimates. Considerable time and effort was made to reduce non-sampling errors in the survey. Quality assurance measures were implemented at each step of the data collection and processing cycle to monitor the quality of the data. These measures included the use of highly skilled interviewers, extensive training of interviewers with respect to the survey procedures and questionnaire, observation of interviewers to detect problems of questionnaire design or misunderstanding of instructions, procedures to ensure that data capture errors were minimized and coding and edit quality checks to verify the processing logic.

8.2.1 The Frame

As the Survey of Self-employment was a supplement to the LFS, the sample frame was the frame used by the LFS.

The SSE sample included individuals who met the SSE criteria at the time of their April LFS interview. Those who did not respond to the LFS and those for whom there was no information about the number of hours they work or their school enrollment were not to be interviewed.

The critical factor for the quality of the SSE sample was the ability of the LFS to identify correctly all respondents who were self-employed in the reference week (whether at work that week or not). As the LFS interviews are conducted both non-proxy and proxy, the SSE interviews (all non-proxy) had to verify the self-employed status of those respondents for whom the LFS information was obtained from a household member. Of the 1,480 respondents asked to confirm their status, only 24 did not consider themselves self-employed in the reference week and were excluded from the survey. On the other hand, due to differences between unedited and edited LFS records, the SSE file includes 21 cases not meeting the SSE eligibility criteria according to the edited LFS file. These records were not removed from the file because during the SSE interview respondents considered themselves self-employed. Overall, the chances of inclusion in the SSE sample of someone who in reality was not self-employed were slim. However, the chances of missing individuals who should be in the SSE sample cannot be assessed easily. The LFS showed an unusually large decline in the number of self-employed between March and April 2000. Most of this decline occurred among the unincorporated self-employed without paid help, especially in construction, other services, finance, and in agriculture. Although in the following months (with exception of May) self-employment registered further decreases, it is possible that the April drop was to some extent caused by the presence of the supplement which made the interviewers apply the definition of self-employed status more rigorously.

8.2.2 Data Collection

Interviewers received a detailed SSE manual and a training workbook with practice cases. Senior interviewers were available to answer interviewers' questions and to try obtaining cooperation of respondents who refused to answer the SSE.

The interviewers who participated in a debriefing after data collection had ended, did not report major difficulties with the survey instrument or the application, but found the requirement of non-proxy interview the biggest challenge of this survey.

During the data processing of the SSE, it became evident that as a result of interviewers' error 28 percent of eligible respondents were not asked the Work Arrangements block of questions that applied to them (those questions were not intended for farmers and fishers). In December 2000, respondents who missed this block were contacted and 90 percent of them answered the missing questions. Those who did not respond have the "not stated" code in place of missing answers.

8.2.3 Data Processing

The computer application had build-in functions preventing question flow errors and out-of range values. Thus, most of the differences between the pre-edited file and the edited file resulted from recoding of "other, specify".

Additionally, a review of answers to the question about work location (WA_Q1A) in the context of respondent's detailed occupation and industry description showed that for some respondents the distinction between "In an office or work space provided to you by your clients" and "In various clients' locations" was not clear. About 40 responses were recoded based on industry and occupation description and answers concerning supplies provided by the clients.

8.2.4 Non-response

Total non-response

A major source of non-sampling errors in surveys is the effect of non-response on the survey results. The extent of non-response varies from partial non-response (failure to answer just one or some questions) to total non-response. Total non-response occurred because the interviewer was either unable to contact the respondent or the respondent refused to participate in the survey. Total non-response was handled by adjusting the weight of households who responded to the survey to compensate for those who did not respond.

The level of non-response to the SSE (39%) was higher than expected. However, the impact of non-response on survey estimates depends not only on the magnitude of non-response, but also on differences between respondents and non-respondents with respect to variables of importance to the survey.

Because the SSE was conducted as a supplement to the LFS, we have access to socio-demographic characteristics of non-respondents and we could compare the two groups. Based on these comparisons, we have concluded that there is no evidence of a serious non-response bias. It appears that the only factor strongly affecting non-response was proxy or non-proxy method of the LFS interview. However, as the interview method is not related to the subject matter of the survey, the possibility of a bias resulting from

significantly higher response to the SSE among persons interviewed non-proxy by the LFS is rather remote.

Partial non-response

In most cases, partial non-response to the survey occurred when the respondent did not understand or misinterpreted a question, refused to answer a question, or could not recall the requested information.

The highest incidence of partial (item) non-response to the SSE occurred in the income set of questions. Of the 3,541 respondents asked about their income, 373 (10 percent) did not answer.

As mentioned before (8.2.2 Data Collection), the Work Arrangements block of questions, due to interviewers' error was missed during over 900 interviews and had to be asked later during the re-interview. Because not all the SSE respondents could be re-interviewed and because of change in the question flow due to recoding "other, specify" answers to WA_Q1A , the number of not answered questions in this section is relatively high (it varies from 97 to 124 cases).

The survey questionnaire contains two hypothetical questions. Question MS_Q7 asks about accepting a paid job in place of self-employment, while question FN_Q15 assesses interest in an insurance program for self-employed. Both have a rather high non-response (respectively, 140 and 221 cases) indicating that for some respondents the hypothetical scenarios were not detailed enough to evoke an answer.

Coefficient of Variation - a measure of sampling error

Since it is an unavoidable fact that estimates from a sample survey are subject to sampling error, sound statistical practice calls for researchers to provide users with some indication of the magnitude of this sampling error. This section of the documentation outlines the measures of sampling error which Statistics Canada commonly uses and which it urges users producing estimates from this microdata file to use also.

The basis for measuring the potential size of sampling errors is the standard error of the estimates derived from survey results. However, because of the large variety of estimates that can be produced from a survey, the standard error of an estimate is usually expressed relative to the estimate to which it pertains. This resulting measure, known as the coefficient of variation (c.v.) of an estimate, is obtained by dividing the standard error of the estimate by the estimate itself and is expressed as a percentage of the estimate.

For example, suppose that, based upon the survey results, one estimates that 42.4% of self-employed Canadians were covered by a health plan, other than provincial medicare, and this estimate is found to have standard error of 0.0103. Then the coefficient of variation of the estimate is calculated as:

$$\left(\frac{0.0103}{0.424} \right) \times 100\% = 2.4\%$$

For more information on calculating approximate coefficients of variation, see section 10.

9.0 Guidelines for Tabulation, Analysis and Release

This section of the documentation outlines the guidelines to be adhered to by users tabulating, analysing, publishing or otherwise releasing any data derived from the survey microdata files. With the aid of these guidelines, users of microdata should be able to produce the same figures as those produced by Statistics Canada and, at the same time, will be able to develop currently unpublished figures in a manner consistent with these established guidelines.

9.1 Rounding Guidelines

In order that estimates for publication or other release derived from these microdata files correspond to those produced by Statistics Canada, users are urged to adhere to the following guidelines regarding the rounding of such estimates:

- a) Estimates in the main body of a statistical table are to be rounded to the nearest hundred units using the normal rounding technique. In normal rounding, if the first or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is raised by one. For example, in normal rounding to the nearest 100, if the last two digits are between 00 and 49, they are changed to 00 and the preceding digit (the hundreds digit) is left unchanged. If the last digits are between 50 and 99 they are changed to 00 and the preceding digit is incremented by 1.
- b) Marginal sub-totals and totals in statistical tables are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units using normal rounding.
- c) Averages, proportions, rates and percentages are to be computed from unrounded components (i.e. numerators and/or denominators) and then are to be rounded themselves to one decimal using normal rounding. In normal rounding to a single digit, if the final or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is increased by 1.
- d) Sums and differences of aggregates (or ratios) are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units (or the nearest one decimal) using normal rounding.
- e) In instances where, due to technical or other limitations, a rounding technique other than normal rounding is used resulting in estimates to be published or otherwise released which differ from corresponding estimates published by Statistics Canada, users are urged to note the reason for such differences in the publication or release document(s).
- f) Under no circumstances are unrounded estimates to be published or otherwise released by users. Unrounded estimates imply greater precision than actually exists.

9.2 Sample Weighting Guidelines for Tabulation

The sample design used for the Survey of Self-employment was not self-weighting. When producing simple estimates, including the production of ordinary statistical tables, users must apply the proper sampling weight.

If proper weights are not used, the estimates derived from the microdata files cannot be considered to be representative of the survey population, and will not correspond to those produced by Statistics Canada.

Users should also note that some software packages may not allow the generation of estimates that exactly match those available from Statistics Canada, because of their treatment of the weight field.

9.2.1 Definitions of types of estimates: Categorical vs. Quantitative

Before discussing how the Survey of Self-employment data can be tabulated and analysed, it is useful to describe the two main types of point estimates of population characteristics which can be generated from the microdata file for the SSE.

Categorical Estimates

Categorical estimates are estimates of the number, or percentage of the surveyed population possessing certain characteristics or falling into some defined category. The number of self-employed who are covered by a health plan, other than provincial medicare or who borrowed money from family or friends are examples of such estimates. An estimate of the number of persons possessing a certain characteristic may also be referred to as an estimate of an aggregate.

Examples of Categorical Questions:

Q: Are you covered by a health plan, other than provincial medicare?
R: Yes / No

Q: From whom did you borrow?
R: From a financial institution/from family or friends/other

Quantitative Estimates

Quantitative estimates are estimates of totals or of means, medians and other measures of central tendency of quantities based upon some or all of the members of the surveyed population. They also specifically involve estimates of the form \hat{X}/\hat{Y} where \hat{X} is an estimate of surveyed population quantity total and \hat{Y} is an estimate of the number of persons in the surveyed population contributing to that total quantity.

An example of a quantitative estimate is the average number of years that passed since the start of current self-employment. The numerator is an estimate of the total number of years respondents have been self-employed, and its denominator is the number of respondents.

Examples of Quantitative Variables :

Number of years respondent has been self-employed (derived from the start year variable).

Age at immigration - derived from birth year and year of immigration

9.2.2 Tabulation of Categorical Estimates

Estimates of the number of people with a certain characteristic can be obtained from the microdata file by summing the final weights of all records possessing the characteristic(s) of interest. Proportions and ratios of the form X/Y are obtained by:

- (a) summing the final weights of records having the characteristic of interest for the numerator (X),
- (b) summing the final weights of records having the characteristic of interest for the denominator (Y), then
- (c) dividing the numerator estimate by the denominator estimate.

9.2.3 Tabulation of Quantitative Estimates

Estimates of quantities can be obtained from the microdata file by multiplying the value of the variable of interest by the final weight for each record, then summing this quantity over all records of interest. The SSE file has only two variables that can be used to create quantitative estimates: number of years in the current self-employment and number of years since immigration. For example, to obtain an estimate of the total number of years in the current self-employment, multiply the value derived from DVSTRYTG (2000 - start year) by the final weight for the record, then sum this value over all records.

To obtain a weighted average of the form X/Y , the numerator (X) is calculated as for a quantitative estimate and the denominator (Y) is calculated as for a categorical estimate. For example, to estimate the average number of years of self-employment,

- (a) estimate the total number of years as described above,
- (b) divide estimate (a) by the estimate of self-employed population (2,078,243).

9.3 Guidelines for Statistical Analysis

The Survey of Self-employment is based upon a complex sample design, with stratification, multiple stages of selection, and unequal probabilities of selection of respondents. Using data from such complex surveys presents problems to analysts because the survey design and the selection probabilities affect the estimation and variance calculation procedures that should be used. In order for survey estimates and analyses to be free from sampling bias, the survey weights must be used.

While many analysis procedures found in statistical packages allow weights to be used, the meaning or definition of the weight in these procedures differ from that which is appropriate in a sample survey framework, with the result that while in many cases the estimates produced by the packages are correct, the variances that are calculated are poor. Variances for simple estimates such as totals, proportions and ratios (for qualitative variables) are provided in the accompanying Sampling Variability Tables.

For other analysis techniques (for example linear regression, logistic regression and analysis of variance), a method exists which can make the variances calculated by the standard packages more meaningful, by incorporating the unequal probabilities of selection. The method rescales the weights so that there is an average weight of 1.

For example, suppose that analysis of all male respondents is required. The steps to rescale the weights are as follows:

- select all respondents from the file who reported SEX=male
- Calculate the AVERAGE weight for these records by summing the original person weights from the microdata file for these records and then dividing by the number of respondents who reported SEX=male
- for each of these respondents, calculate a RESCALED weight equal to the original person weight divided by the AVERAGE weight
- perform the analysis for these respondents using the RESCALED weight.

However, because the stratification and clustering of the sample's design are still not taken into account, the variance estimates calculated in this way are likely to be under-estimates.

The calculation of truly meaningful variance estimates requires detailed knowledge of the design of the survey. Such detail cannot be given in this microdata file because of confidentiality. Variances that take the complete sample design into account can be calculated for many statistics by Statistics Canada on a cost recovery basis.

9.4 C.V. Release Guidelines

Before releasing and/or publishing any estimate from the Survey of Self-employment, users should first determine the quality level of the estimate. The quality levels are *acceptable*, *marginal* and *unacceptable*. Data quality is affected by both sampling and non-sampling errors as discussed in section 8. However for this purpose, the quality level of an estimate will be determined only on the basis of sampling error as reflected by the coefficient of variation as shown in the table below. Nonetheless users should be sure to read section 8 to be more fully aware of the quality characteristics of these data.

First, the number of respondents who contribute to the calculation of the estimate should be determined. If this number is less than 30, the weighted estimate should be considered to be of unacceptable quality.

For weighted estimates based on sample sizes of 30 or more, users should determine the coefficient of variation of the estimate and follow the guidelines below. These quality level guidelines should be applied to weighted rounded estimates.

All estimates can be considered releasable. However, those of marginal or unacceptable quality level must be accompanied by a warning to caution subsequent users.

Quality Level Guidelines

Quality Level of Estimate	Guidelines
1. Acceptable	<p>Estimates have: a sample size of 30 or more, and low coefficients of variation in the range 0.0% - 16.5%</p> <p>No warning is required.</p>
2. Marginal	<p>Estimates have: a sample size of 30 or more, and high coefficients of variation in the range 16.6% - 33.3%.</p> <p>Estimates should be flagged with the letter M (or some similar identifier). They should be accompanied by a warning to caution subsequent users about the high levels of error, associated with the estimates.</p>
3. Unacceptable	<p>Estimates have: a sample size of less than 30, or very high coefficients of variation in excess of 33.3%.</p> <p>Statistics Canada recommends not to release estimates of unacceptable quality. However, if the user chooses to do so then estimates should be flagged with the letter U (or some similar identifier) and the following warning should accompany the estimates:</p> <p>"The user is advised that . . . (specify the data) . . . do not meet Statistics Canada's quality standards for this statistical program. Conclusions based on these data will be unreliable, and most likely invalid. These data and any consequent findings should not be published. If the user chooses to publish these data or findings, then this disclaimer must be published with the data."</p>

10.0 Approximate Sampling Variability Tables

In order to supply coefficients of variation which would be applicable to a wide variety of categorical estimates produced from this microdata file and which could be readily accessed by the user, a set of Approximate Sampling Variability Tables has been produced. These "look-up" tables allow the user to obtain an approximate coefficient of variation based on the size of the estimate calculated from the survey data.

The coefficients of variation are derived using the variance formula for simple random sampling and incorporating a factor which reflects the multi-stage, clustered nature of the sample design. This factor, known as the design effect, was determined by first calculating design effects for a wide range of characteristics and then choosing from among these a conservative value to be used in the look-up tables which would then apply to the entire set of characteristics.

The table below shows the design effects, sample sizes and population counts by region which were used to produce the Approximate Sampling Variability Tables.

PROVINCE	DESIGN EFFECT	SAMPLE SIZE	POPULATION
Atlantic Provinces	1.52	593	116250
Quebec	2.03	724	460335
Ontario	1.82	1162	771490
Prairies	1.82	1120	417188
British Columbia	1.37	416	312980
Canada	2.00	4015	2078243

All coefficients of variation in the Approximate Sampling Variability Tables are approximate and, therefore, unofficial. Estimates of actual variance for specific variables may be obtained from Statistics Canada on a cost-recovery basis. Given that the approximate coefficients of variation are conservative, the use of actual variance estimates would allow users to move from one quality level to another one. For example, an estimate of *marginal* quality could become *acceptable* when based on an actual c.v.

Remember: if the number of observations on which an estimate is based is less than 30, the weighted estimate should not be released regardless of the value of the coefficient of variation for this estimate. This is because the formulas used for estimating the variance do not hold true for small sample sizes.

10.1 How to use the C.V. tables for Categorical Estimates

The following rules should enable the user to determine the approximate coefficients of variation from the Sampling Variability Tables for estimates of the number, proportion or percentage of the surveyed population possessing a certain characteristic and for ratios and differences between such estimates.

Rule 1: Estimates of Numbers Possessing a Characteristic (Aggregates)

The coefficient of variation depends only on the size of the estimate itself. On the Sampling Variability Table for the appropriate geographic area, locate the estimated number in the left-most column of the table (headed "Numerator of Percentage") and follow the asterisks (if any) across to the first figure encountered. This figure is the approximate coefficient of variation.

Rule 2: Estimates of Proportions or Percentages Possessing a Characteristic

The coefficient of variation of an estimated proportion or percentage depends on both the size of the proportion or percentage and the size of the total upon which the proportion or percentage is based. Estimated proportions or percentages are relatively more reliable than the corresponding estimates of the numerator of the proportion or percentage, when the proportion or percentage is based upon a sub-group of the population. For example, the proportion of "self-employed who are covered by dental insurance thorough spouse" is more reliable than the estimated number of "self-employed who are covered by dental insurance through spouse". (Note that in the tables the coefficients of variation decline in value reading from left to right).

When the proportion or percentage is based upon the total population of the geographic area covered by the table, the c.v. of the proportion or percentage is the same as the c.v. of the numerator of the proportion or percentage. In this case, Rule 1 can be used.

When the proportion or percentage is based upon a subset of the total population (e.g. those in a particular sex or age group), reference should be made to the proportion or percentage (across the top of the table) and to the numerator of the proportion or percentage (down the left side of the table). The intersection of the appropriate row and column gives the coefficient of variation.

Rule 3: Estimates of Differences Between Aggregates or Percentages

The standard error of a difference between two estimates is approximately equal to the square root of the sum of squares of each standard error considered separately. That is, the standard error of a difference ($\hat{d} = \hat{X}_1 - \hat{X}_2$) is:

$$\sigma_{\hat{d}} = \sqrt{(\hat{X}_1 \alpha_1)^2 + (\hat{X}_2 \alpha_2)^2}$$

where \hat{X}_1 is estimate 1, \hat{X}_2 is estimate 2, and α_1 and α_2 are the coefficients of variation of \hat{X}_1 and \hat{X}_2 respectively. The coefficient of variation of \hat{d} is given by $\sigma_{\hat{d}}/\hat{d}$. This formula is accurate for the difference between separate and uncorrelated characteristics, but is only approximate otherwise.

Rule 4: Estimates of Ratios

In the case where the numerator is a subset of the denominator, the ratio should be converted to a percentage and Rule 2 applied. This would apply, for example, to the case where the denominator is "the number of self-employed with dental insurance" and the numerator is "the number of self-employed with dental insurance through spouse".

In the case where the numerator is not a subset of the denominator, as for example, the ratio of the number of "self-employed with dental insurance" compared to the number of "self-employed

with health insurance”, the standard deviation of the ratio of the estimates is approximately equal to the square root of the sum of squares of each coefficient of variation considered separately multiplied by R. That is, the standard error of a ratio ($\hat{R} = X_1 / X_2$) is:

$$\sigma_{\hat{R}} = \hat{R} \sqrt{\alpha_1^2 + \alpha_2^2}$$

where α_1 and α_2 are the coefficients of variation of X_1 and X_2 respectively. The coefficient of variation of R is given by σ_R/R . The formula will tend to overstate the error, if X_1 and X_2 are positively correlated and understate the error if X_1 and X_2 are negatively correlated.

Rule 5: Estimates of Differences of Ratios

In this case, Rules 3 and 4 are combined. The coefficients of variation for the two ratios are first determined using Rule 4, and then the c.v. of their difference is found using Rule 3.

10.1.1 Examples of using the C.V. tables for Categorical Estimates

The following 'real life' examples are included to assist users in applying the foregoing rules.

Example 1 : Estimates of Numbers Possessing a Characteristic (Aggregates)

Suppose that a user estimates that 881,025 self-employed were covered by a health plan, other than provincial medicare . How does the user determine the coefficient of variation of this estimate?

- (1) Refer to the c.v. table for CANADA.
- (2) The estimated aggregate (881,025) does not appear in the left-hand column (the 'Numerator of Percentage' column), so it is necessary to use the figure closest to it, namely 1,000,000.
- (3) The coefficient of variation for an estimated aggregate is found by referring to the first non-asterisk entry on that row, namely, 2.3%.
- (4) So the approximate coefficient of variation of the estimate is 2.3%.The finding that there were 881,025 self-employed who were covered by health insurance is publishable with no qualifications.

Example 2 : Estimates of Proportions or Percentages Possessing a Characteristic

Suppose that the user estimates that 389,305/881,025=44.2% of self-employed who were covered by health insurance had the insurance through spouse. How does the user determine the coefficient of variation of this estimate?

- (1) Refer to the table for CANADA.

- (2) Because the estimate is a percentage which is based on a subset of the total population (i.e., self-employed with an insurance), it is necessary to use both the percentage (44.2%) and the numerator portion of the percentage (389,305) in determining the coefficient of variation.
- (3) The numerator, 389,305, does not appear in the left-hand column (the 'Numerator of Percentage' column) so it is necessary to use the figure closest to it, namely 400,000. Similarly, the percentage estimate does not appear as any of the column headings, so it is necessary to use the figure closest to it, 40.0%.
- (4) The figure at the intersection of the row and column used, namely 3.9% is the coefficient of variation to be used.
- (5) So the approximate coefficient of variation of the estimate is 3.9%. The finding that 44.2% of self-employed with health insurance had the insurance through spouse can be published with no qualifications.

Example 3 : Estimates of Differences Between Aggregates or Percentages

Suppose that a user estimates that $204,746/666,863=30.7\%$ of female self-employed took formal training in the reference year, while $354,940/1,411,380=25.1\%$ of male self-employed took formal training. How does the user determine the coefficient of variation of the difference between these two estimates?

- (1) Using the c.v. table for CANADA in the same manner as described in example 2 gives the c.v. of the estimate for females as 6.0%, and the c.v. of the estimate for males as 4.7%.

Using rule 3, the standard error of a difference ($\hat{d} = \bar{X}_1 - \bar{X}_2$) is:

$$\sigma_{\hat{d}} = \sqrt{(\hat{X}_1 \alpha_1)^2 + (\hat{X}_2 \alpha_2)^2}$$

where \bar{X}_1 is estimate 1, \bar{X}_2 is estimate 2, and α_1 and α_2 are the coefficients of variation of \bar{X}_1 and \bar{X}_2 respectively.

That is, the standard error of the difference $\hat{d} = (0.307-0.251) = 0.056$ is:

$$\begin{aligned} \sigma_{\hat{d}} &= \sqrt{[(0.307)(0.06)]^2 + [(0.251)(0.047)]^2} \\ &= \sqrt{(0.000339) + (0.00139)} \\ &= 0.0219 \end{aligned}$$

- (3) The coefficient of variation of \hat{d} is given by $\sigma_{\hat{d}}/\hat{d} = 0.0219/0.056 = 0.391$.

- (4) So the approximate coefficient of variation of the difference between the estimates is 39%. This estimate should not be released. However, if the user chooses to do so then estimates should be flagged and accompanied by the warning, see Quality Level Guidelines.

Example 4 : Estimates of Ratios

Suppose that the user estimates that 204,746 female self-employed took formal training in the reference year, while 354,940 of male self-employed took formal training. The user is interested in comparing the estimate of women versus that of men in the form of a ratio. How does the user determine the coefficient of variation of this estimate?

- (1) First of all, this estimate is a ratio estimate, where the numerator of the estimate ($= X_1$) is the number of female self-employed who took formal training. The denominator of the estimate ($= X_2$) is the number of male self-employed who took formal training.
 - (2) Refer to the table for CANADA.
 - (3) The numerator of this ratio estimate is 204,746. The figure closest to it is 200,000. The coefficient of variation for this estimate is found by referring to the first non-asterisk entry on that row, namely, 6.8%.
 - (4) The denominator of this ratio estimate is 354,940. The figure closest to it is 350,000. The coefficient of variation for this estimate is found by referring to the first non-asterisk entry on that row, namely, 4.9%.
- (5) So the approximate coefficient of variation of the ratio estimate is given by rule 4, which is,

$$\alpha_{\hat{R}} = \sqrt{\alpha_1^2 + \alpha_2^2}$$

where α_1 and α_2 are the coefficients of variation of X_1 and X_2 respectively.

That is ,

$$\begin{aligned} \alpha_{\hat{R}} &= \sqrt{(0.068)^2 + (0.049)^2} \\ &= 0.0838 \end{aligned}$$

The obtained ratio of self-employed women who took formal training versus men is 204,746/354,940 which is 0.58. The coefficient of variation of this estimate is 8.4%, which is releasable with no qualifications.

10.2 How to use the C.V. tables to obtain Confidence Limits

Although coefficients of variation are widely used, a more intuitively meaningful measure of sampling error is the confidence interval of an estimate. A confidence interval constitutes a statement on the level of confidence that the true value for the population lies within a specified range of values. For example a 95% confidence interval can be described as follows:

If sampling of the population is repeated indefinitely, each sample leading to a new confidence interval for an estimate, then in 95% of the samples the interval will cover the true population value.

Using the standard error of an estimate, confidence intervals for estimates may be obtained under the assumption that under repeated sampling of the population, the various estimates obtained for a population characteristic are normally distributed about the true population value. Under this assumption, the chances are about 68 out of 100 that the difference between a sample estimate and the true population value would be less than one standard error, about 95 out of 100 that the difference would be less than two standard errors, and about 99 out of 100 that the differences would be less than three standard errors. These different degrees of confidence are referred to as the confidence levels.

Confidence intervals for an estimate, \hat{X} , are generally expressed as two numbers, one below the estimate and one above the estimate, as $(\hat{X}-k, \hat{X}+k)$ where k is determined depending upon the level of confidence desired and the sampling error of the estimate.

Confidence intervals for an estimate can be calculated directly from the Approximate Sampling Variability Tables by first determining from the appropriate table the coefficient of variation of the estimate \hat{X} , and then using the following formula to convert to a confidence interval CI:

$$CI_X = [\hat{X} - t\hat{X}\alpha_{\hat{X}}, \hat{X} + t\hat{X}\alpha_{\hat{X}}]$$

where α_X is the determined coefficient of variation of \hat{X} , and

- t = 1 if a 68% confidence interval is desired
- t = 1.6 if a 90% confidence interval is desired
- t = 2 if a 95% confidence interval is desired
- t = 3 if a 99% confidence interval is desired.

Note: Release guidelines which apply to the estimate also apply to the confidence interval. For example, if the estimate is unacceptable, then the confidence interval is also unacceptable.

10.2.1 Example of using the C.V. tables to obtain confidence limits

A 95% confidence interval for the estimated proportion of self-employed who had health insurance through their spouse (from Example 2, section 10.1.1) would be calculated as follows.

$$\hat{X} = 44.2\% \text{ (or expressed as a proportion} = 0.442)$$

$$t = 2$$

$$\alpha_X = 3.9\% \text{ (0.039 expressed as a proportion) is the coefficient of variation of this estimate as determined from the tables.}$$

$$CI_x = \{0.442 - (2) (0.442) (0.039), 0.442 + (2) (0.442) (0.039)\}$$

$$CI_x = \{0.442 - 0.034, 0.442 + 0.034\}$$

$$CI_x = \{0.408, 0.476\}$$

With 95% confidence it can be said that between 40.8% and 47.6% of self-employed that had health insurance (other than the provincial one), had it through their spouse.

10.3 How to use the C.V. tables to do a t-test

Standard errors may also be used to perform hypothesis testing, a procedure for distinguishing between population parameters using sample estimates. The sample estimates can be numbers, averages, percentages, ratios, etc. Tests may be performed at various levels of significance, where a level of significance is the probability of concluding that the characteristics are different when, in fact, they are identical.

Let X_1 and X_2 be sample estimates for 2 characteristics of interest. Let the standard error on the difference $X_1 - X_2$ be σ_d .

If $t = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_d}$ is between -2 and 2, then no conclusion about the difference between

the characteristics is justified at the 5% level of significance. If however, this ratio is smaller than -2 or larger than +2, the observed difference is significant at the 0.05 level. That is to say that the characteristics are significant.

10.3.1 Example of using the C.V. tables to do a t-test

Let us suppose we wish to test, at 5% level of significance, the hypothesis that there is no difference between the proportion of female self-employed reporting having taken a formal training in the reference year and the proportion of male self-employed reporting having taken a formal training. From example 3, section 10.1.1, the standard error of the difference between these two estimates was found to be = .0219. Hence ,

$$t = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_d} = \frac{0.307 - 0.251}{0.0219} = \frac{0.056}{0.0219} = 2.56.$$

Since $t = 2.56$ is larger than 2, it must be concluded that there is a significant difference between the two estimates at the 0.05 level of significance. In other words, the proportion of women who took formal training is larger, than such a proportion among men.

10.4 Coefficients of Variation for Quantitative Estimates

For quantitative estimates, special tables would have to be produced to determine their sampling error. Since most of the variables for the Survey of Self-employment are primarily categorical in nature, this has not been done.

As a general rule, however, the coefficient of variation of a quantitative total will be larger than the coefficient of variation of the corresponding category estimate (i.e., the estimate of the number of persons contributing to the quantitative estimate). If the corresponding category estimate is not releasable, the quantitative estimate will not be either. For example, the coefficient of variation of the total number of years since immigrating to Canada will be larger than the coefficient of variation of the corresponding proportion of self-employed who immigrated to Canada). Hence if the coefficient of variation of the proportion is unacceptable, then the coefficient of variation of the corresponding quantitative estimate will also be unacceptable.

Coefficients of variation of such estimates can be derived as required for a specific estimate using a technique known as pseudo replication. This involves dividing the records on the microdata files into subgroups (or replicates) and determining the variation in the estimate from replicate to replicate. Users wishing to derive coefficients of variation for quantitative estimates may contact Statistics Canada for advice on the allocation of records to appropriate replicates and the formulae to be used in these calculations.

10.5 Release cut-off's for the Survey of Self-employment

The minimum size of the estimate at the provincial, regional and Canada levels are specified in the table below. Estimates smaller than the minimum size given in the "Acceptable" column should not be released without a warning (see Quality Level Guidelines).

Table of Release Cut-offs

Province	Acceptable	Marginal	Unacceptable
Atlantic Provinces	10,000+	3,000 - 9,999	under 3,000
Quebec	43,000+	12,000 - 42,999	under 12,000
Ontario	42,000+	11,000 - 41,999	under 11,000
Prairie Provinces	24,000+	6,000 - 23,999	under 6,000
British Columbia	34,000+	9,000 - 33,999	under 9,000
Newfoundland and Labrador	9,000+	3,000 - 8,999	under 3,000
Prince Edward Island	3,000+	1,000 - 2,999	under 1,000
Nova Scotia	10,000+	3,000 - 9,999	under 3,000
New Brunswick	9,000+	3,000 - 8,999	under 3,000
Manitoba	13,000+	4,000 - 12,999	under 4,000
Saskatchewan	12,000+	3,000 - 11,999	under 3,000
Alberta	30,000+	8,000 - 29,999	under 8,000
CANADA	38,000+	10,000 - 37,999	under 10,000

10.6 C.V. Tables

2000 Survey of Self-employment

Approximate Sampling Variability Tables for Atlantic provinces

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	*****	54.2	53.9	53.1	51.7	50.2	48.7	47.2	45.6	43.9	42.2	38.5	29.8	17.2
2	*****		38.1	37.5	36.5	35.5	34.4	33.3	32.2	31.0	29.8	27.2	21.1	12.2
3	*****			30.6	29.8	29.0	28.1	27.2	26.3	25.3	24.3	22.2	17.2	9.9
4	*****			26.5	25.8	25.1	24.3	23.6	22.8	21.9	21.1	19.3	14.9	8.6
5	*****			23.7	23.1	22.4	21.8	21.1	20.4	19.6	18.9	17.2	13.3	7.7
6	*****				21.1	20.5	19.9	19.3	18.6	17.9	17.2	15.7	12.2	7.0
7	*****				19.5	19.0	18.4	17.8	17.2	16.6	15.9	14.6	11.3	6.5
8	*****				18.3	17.7	17.2	16.7	16.1	15.5	14.9	13.6	10.5	6.1
9	*****				17.2	16.7	16.2	15.7	15.2	14.6	14.1	12.8	9.9	5.7
10	*****				16.3	15.9	15.4	14.9	14.4	13.9	13.3	12.2	9.4	5.4
11	*****				15.6	15.1	14.7	14.2	13.7	13.2	12.7	11.6	9.0	5.2
12	*****					14.5	14.1	13.6	13.2	12.7	12.2	11.1	8.6	5.0
13	*****					13.9	13.5	13.1	12.6	12.2	11.7	10.7	8.3	4.8
14	*****					13.4	13.0	12.6	12.2	11.7	11.3	10.3	8.0	4.6
15	*****					13.0	12.6	12.2	11.8	11.3	10.9	9.9	7.7	4.4
16	*****					12.5	12.2	11.8	11.4	11.0	10.5	9.6	7.5	4.3
17	*****					12.2	11.8	11.4	11.0	10.6	10.2	9.3	7.2	4.2
18	*****						11.5	11.1	10.7	10.3	9.9	9.1	7.0	4.1
19	*****						11.2	10.8	10.5	10.1	9.7	8.8	6.8	4.0
20	*****						10.9	10.5	10.2	9.8	9.4	8.6	6.7	3.9
21	*****						10.6	10.3	9.9	9.6	9.2	8.4	6.5	3.8
22	*****						10.4	10.1	9.7	9.4	9.0	8.2	6.4	3.7
23	*****						10.2	9.8	9.5	9.2	8.8	8.0	6.2	3.6
24	*****							9.6	9.3	9.0	8.6	7.9	6.1	3.5
25	*****							9.4	9.1	8.8	8.4	7.7	6.0	3.4
30	*****								8.3	8.0	7.7	7.0	5.4	3.1
35	*****									7.4	7.1	6.5	5.0	2.9
40	*****										6.9	6.7	6.1	4.7
45	*****											6.3	5.7	4.4
50	*****												5.4	4.2
55	*****													5.2
60	*****													4.0
65	*****													3.9
70	*****													3.7
75	*****													3.6
80	*****													3.4
85	*****													3.3
90	*****													3.3
95	*****													1.9
100	*****													1.9

NOTE: FOR CORRECT USAGE OF THESE TABLES PLEASE REFER TO MICRODATA DOCUMENTATION

2000 Survey of Self-employment – User Guide

2000 Survey of Self-employment

Approximate Sampling Variability Tables for Quebec

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	*****	113.0	112.4	110.6	107.7	104.7	101.5	98.3	95.0	91.5	87.9	80.3	62.2	35.9
2	*****	79.9	79.5	78.2	76.2	74.0	71.8	69.5	67.2	64.7	62.2	56.8	44.0	25.4
3	*****	65.2	64.9	63.9	62.2	60.4	58.6	56.8	54.8	52.8	50.8	46.3	35.9	20.7
4	*****	56.5	56.2	55.3	53.8	52.3	50.8	49.2	47.5	45.8	44.0	40.1	31.1	17.9
5	*****	50.3	49.5	48.2	46.8	45.4	44.0	42.5	40.9	39.3	35.9	27.8	16.1	
6	*****	45.9	45.2	44.0	42.7	41.5	40.1	38.8	37.4	35.9	32.8	25.4	14.7	
7	*****	42.5	41.8	40.7	39.6	38.4	37.2	35.9	34.6	33.2	30.3	23.5	13.6	
8	*****	39.7	39.1	38.1	37.0	35.9	34.8	33.6	32.4	31.1	28.4	22.0	12.7	
9	*****	37.5	36.9	35.9	34.9	33.8	32.8	31.7	30.5	29.3	26.8	20.7	12.0	
10	*****	35.0	34.1	33.1	32.1	31.1	30.0	28.9	27.8	25.4	19.7	11.4		
11	*****	33.4	32.5	31.6	30.6	29.6	28.6	27.6	26.5	24.2	18.7	10.8		
12	*****	31.9	31.1	30.2	29.3	28.4	27.4	26.4	25.4	23.2	17.9	10.4		
13	*****	30.7	29.9	29.0	28.2	27.3	26.3	25.4	24.4	22.3	17.2	10.0		
14	*****	29.6	28.8	28.0	27.1	26.3	25.4	24.5	23.5	21.5	16.6	9.6		
15	*****	28.6	27.8	27.0	26.2	25.4	24.5	23.6	22.7	20.7	16.1	9.3		
16	*****	27.7	26.9	26.2	25.4	24.6	23.7	22.9	22.0	20.1	15.5	9.0		
17	*****	26.8	26.1	25.4	24.6	23.8	23.0	22.2	21.3	19.5	15.1	8.7		
18	*****	26.1	25.4	24.7	23.9	23.2	22.4	21.6	20.7	18.9	14.7	8.5		
19	*****	25.4	24.7	24.0	23.3	22.6	21.8	21.0	20.2	18.4	14.3	8.2		
20	*****	24.7	24.1	23.4	22.7	22.0	21.2	20.5	19.7	17.9	13.9	8.0		
21	*****	24.1	23.5	22.8	22.2	21.5	20.7	20.0	19.2	17.5	13.6	7.8		
22	*****	23.6	23.0	22.3	21.6	21.0	20.2	19.5	18.7	17.1	13.3	7.7		
23	*****	23.1	22.5	21.8	21.2	20.5	19.8	19.1	18.3	16.7	13.0	7.5		
24	*****	22.0	21.4	20.7	20.1	19.4	18.7	17.9	17.1	16.4	12.7	7.3		
25	*****	21.5	20.9	20.3	19.7	19.0	18.3	17.6	16.9	16.1	12.4	7.2		
30	*****	19.7	19.1	18.5	17.9	17.3	16.7	16.1	15.5	14.7	11.4	6.6		
35	*****	18.2	17.7	17.2	16.6	16.1	15.5	14.9	14.3	13.6	10.5	6.1		
40	*****	17.0	16.5	16.1	15.5	15.0	14.5	13.9	13.3	12.7	9.8	5.7		
45	*****	16.1	15.6	15.1	14.7	14.2	13.6	13.1	12.5	12.0	9.3	5.4		
50	*****	14.8	14.4	13.9	13.4	12.9	12.4	11.9	11.4	10.9	8.8	5.1		
55	*****	14.1	13.7	13.3	12.8	12.3	11.8	11.3	10.8	10.3	8.4	4.8		
60	*****	13.5	13.1	12.7	12.3	11.8	11.4	10.9	10.4	9.9	8.0	4.6		
65	*****	13.0	12.6	12.2	11.8	11.4	10.9	10.4	9.9	9.4	7.7	4.5		
70	*****	12.1	11.8	11.4	11.0	10.6	10.2	9.7	9.2	8.7	7.4	4.3		
75	*****	11.7	11.4	11.0	10.6	10.2	9.7	9.3	8.8	8.3	7.2	4.1		
80	*****	11.4	11.0	10.6	10.2	9.7	9.3	8.8	8.3	7.8	7.0	4.0		
85	*****	11.0	10.7	10.3	9.9	9.5	9.0	8.6	8.1	7.6	6.7	3.9		
90	*****	10.7	10.4	10.0	9.6	9.2	8.8	8.3	7.8	7.3	6.6	3.8		
95	*****	10.1	9.7	9.4	9.0	8.6	8.1	7.6	7.1	6.6	6.4	3.7		
100	*****	9.8	9.5	9.2	8.8	8.4	7.9	7.4	6.9	6.4	6.2	3.6		
125	*****	8.5	8.2	7.9	7.5	7.1	6.7	6.3	5.9	5.5	5.6	3.2		
150	*****	7.5	7.2	6.8	6.4	6.0	5.6	5.2	4.8	4.4	5.1	2.9		
200	*****	5.7	5.4	5.0	4.6	4.2	3.8	3.4	3.0	2.6	4.4	2.5		
250	*****	3.9	3.6	3.2	2.8	2.4	2.0	1.6	1.2	0.8	3.9	2.3		
300	*****	3.6	3.2	2.8	2.4	2.0	1.6	1.2	0.8	0.4	3.6	2.1		
350	*****	1.9	1.8	1.6	1.4	1.2	1.0	0.8	0.6	0.4	1.9	1.9		
400	*****	1.8	1.6	1.4	1.2	1.0	0.8	0.6	0.4	0.2	1.8	1.8		

NOTE: FOR CORRECT USAGE OF THESE TABLES PLEASE REFER TO MICRODATA DOCUMENTATION

2000 Survey of Self-employment – User Guide

2000 Survey of Self-employment

Approximate Sampling Variability Tables for Ontario

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	*****	109.3	108.7	107.1	104.2	101.3	98.2	95.1	91.9	88.6	85.1	77.7	60.2	34.7
2	*****	77.3	76.9	75.7	73.7	71.6	69.5	67.3	65.0	62.6	60.2	54.9	42.5	24.6
3	*****	63.1	62.8	61.8	60.2	58.5	56.7	54.9	53.1	51.1	49.1	44.8	34.7	20.1
4	*****	54.6	54.4	53.5	52.1	50.6	49.1	47.6	46.0	44.3	42.5	38.8	30.1	17.4
5	*****	48.9	48.6	47.9	46.6	45.3	43.9	42.5	41.1	39.6	38.1	34.7	26.9	15.5
6	*****	44.6	44.4	43.7	42.5	41.3	40.1	38.8	37.5	36.2	34.7	31.7	24.6	14.2
7	*****	41.3	41.1	40.5	39.4	38.3	37.1	36.0	34.7	33.5	32.2	29.4	22.7	13.1
8	*****	38.4	37.9	36.8	35.8	34.7	33.6	32.5	31.3	30.1	27.5	21.3	12.3	7.4
9	*****	36.2	35.7	34.7	33.8	32.7	31.7	30.6	29.5	28.4	25.9	20.1	11.6	6.5
10	*****	34.4	33.9	33.0	32.0	31.1	30.1	29.1	28.0	26.9	24.6	19.0	11.0	6.0
11	*****	32.8	32.3	31.4	30.5	29.6	28.7	27.7	26.7	25.7	23.4	18.1	10.5	5.5
12	*****	31.4	30.9	30.1	29.2	28.4	27.5	26.5	25.6	24.6	22.4	17.4	10.0	5.0
13	*****	30.2	29.7	28.9	28.1	27.2	26.4	25.5	24.6	23.6	21.5	16.7	9.6	4.6
14	*****	29.1	28.6	27.9	27.1	26.3	25.4	24.6	23.7	22.7	20.8	16.1	9.3	4.3
15	*****	28.1	27.6	26.9	26.1	25.4	24.6	23.7	22.9	22.0	20.1	15.5	9.0	4.0
16	*****	26.8	26.1	25.3	24.6	23.8	23.0	22.1	21.3	19.4	15.0	8.7	3.7	3.7
17	*****	26.0	25.3	24.6	23.8	23.1	22.3	21.5	20.6	18.8	14.6	8.4	3.4	3.4
18	*****	25.2	24.6	23.9	23.2	22.4	21.7	20.9	20.1	18.3	14.2	8.2	3.2	3.2
19	*****	24.6	23.9	23.2	22.5	21.8	21.1	20.3	19.5	17.8	13.8	8.0	3.0	3.0
20	*****	23.9	23.3	22.6	22.0	21.3	20.5	19.8	19.0	17.4	13.5	7.8	2.8	2.8
21	*****	23.4	22.7	22.1	21.4	20.8	20.1	19.3	18.6	16.9	13.1	7.6	2.6	2.6
22	*****	22.8	22.2	21.6	20.9	20.3	19.6	18.9	18.1	16.6	12.8	7.4	2.4	2.4
23	*****	22.3	21.7	21.1	20.5	19.8	19.2	18.5	17.7	16.2	12.5	7.2	2.2	2.2
24	*****	21.9	21.3	20.7	20.1	19.4	18.8	18.1	17.4	15.9	12.3	7.1	2.1	2.1
25	*****	21.4	20.8	20.3	19.6	19.0	18.4	17.7	17.0	15.5	12.0	6.9	2.0	2.0
30	*****	19.5	19.0	18.5	17.9	17.4	16.8	16.2	15.5	14.2	11.0	6.3	1.8	1.8
35	*****	18.1	17.6	17.1	16.6	16.1	15.5	15.0	14.4	13.1	10.2	5.9	1.6	1.6
40	*****	16.5	16.0	15.5	15.0	14.5	14.0	13.5	13.0	11.6	9.5	5.5	1.5	1.5
45	*****	15.5	15.1	14.6	14.2	13.7	13.2	12.7	12.2	10.8	9.0	5.2	1.4	1.4
50	*****	14.7	14.3	13.9	13.5	13.0	12.5	12.0	11.5	10.0	8.5	4.9	1.3	1.3
55	*****	14.1	13.7	13.2	12.8	12.4	11.9	11.5	11.0	9.5	8.1	4.7	1.2	1.2
60	*****	13.5	13.1	12.7	12.3	11.9	11.4	11.0	10.5	9.0	7.8	4.5	1.1	1.1
65	*****	12.9	12.6	12.2	11.8	11.4	11.0	10.6	10.2	8.7	7.5	4.3	1.0	1.0
70	*****	12.5	12.1	11.7	11.4	11.0	10.6	10.2	9.8	8.3	7.2	4.2	0.9	0.9
75	*****	12.0	11.7	11.3	11.0	10.6	10.2	9.8	9.4	8.0	6.9	4.0	0.8	0.8
80	*****	11.3	11.0	10.6	10.3	9.9	9.5	9.1	8.7	7.3	6.3	3.9	0.7	0.7
85	*****	11.0	10.7	10.3	10.0	9.6	9.2	8.8	8.4	7.0	6.0	3.8	0.6	0.6
90	*****	10.7	10.4	10.0	9.7	9.3	9.0	8.6	8.2	6.8	5.8	3.7	0.5	0.5
95	*****	10.4	10.1	9.8	9.4	9.1	8.7	8.3	7.9	6.5	5.5	3.6	0.4	0.4
100	*****	10.1	9.8	9.5	9.2	8.9	8.5	8.1	7.7	6.3	5.3	3.5	0.3	0.3
125	*****	8.8	8.5	8.2	7.9	7.6	7.2	6.9	6.5	5.1	4.2	3.1	0.2	0.2
150	*****	8.0	7.8	7.5	7.2	6.9	6.6	6.3	6.0	4.6	3.7	2.8	0.1	0.1
200	*****	6.5	6.3	6.0	5.7	5.4	5.1	4.8	4.5	3.1	2.2	1.6	0.1	0.1
250	*****	5.6	5.4	5.1	4.8	4.5	4.2	3.9	3.6	2.2	1.3	1.0	0.1	0.1
300	*****	4.9	4.7	4.4	4.1	3.8	3.5	3.2	2.9	1.5	1.1	0.8	0.1	0.1
350	*****	4.2	4.0	3.7	3.4	3.1	2.8	2.5	2.2	1.1	0.8	0.6	0.1	0.1
400	*****	3.0	2.8	2.6	2.4	2.2	2.0	1.8	1.6	0.9	0.6	0.4	0.1	0.1
450	*****	2.8	2.6	2.4	2.2	2.0	1.8	1.6	1.4	0.8	0.5	0.3	0.1	0.1
500	*****	2.7	2.5	2.3	2.1	1.9	1.7	1.5	1.3	0.7	0.4	0.2	0.1	0.1

NOTE: FOR CORRECT USAGE OF THESE TABLES PLEASE REFER TO MICRODATA DOCUMENTATION

2000 Survey of Self-employment – User Guide

2000 Survey of Self-employment

Approximate Sampling Variability Tables for Manitoba

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	*****		62.6	61.7	60.0	58.3	56.6	54.8	52.9	51.0	49.0	44.7	34.6	20.0
2	*****			43.6	42.4	41.2	40.0	38.7	37.4	36.1	34.6	31.6	24.5	14.1
3	*****			35.6	34.6	33.7	32.7	31.6	30.6	29.4	28.3	25.8	20.0	11.5
4	*****				30.0	29.2	28.3	27.4	26.5	25.5	24.5	22.4	17.3	10.0
5	*****				26.8	26.1	25.3	24.5	23.7	22.8	21.9	20.0	15.5	8.9
6	*****				24.5	23.8	23.1	22.4	21.6	20.8	20.0	18.3	14.1	8.2
7	*****				22.7	22.0	21.4	20.7	20.0	19.3	18.5	16.9	13.1	7.6
8	*****					20.6	20.0	19.4	18.7	18.0	17.3	15.8	12.2	7.1
9	*****					19.4	18.9	18.3	17.6	17.0	16.3	14.9	11.5	6.7
10	*****					18.4	17.9	17.3	16.7	16.1	15.5	14.1	11.0	6.3
11	*****					17.6	17.1	16.5	16.0	15.4	14.8	13.5	10.4	6.0
12	*****						16.3	15.8	15.3	14.7	14.1	12.9	10.0	5.8
13	*****						15.7	15.2	14.7	14.1	13.6	12.4	9.6	5.5
14	*****						15.1	14.6	14.1	13.6	13.1	12.0	9.3	5.3
15	*****						14.6	14.1	13.7	13.2	12.7	11.5	8.9	5.2
16	*****							13.7	13.2	12.8	12.2	11.2	8.7	5.0
17	*****							13.3	12.8	12.4	11.9	10.8	8.4	4.9
18	*****							12.9	12.5	12.0	11.5	10.5	8.2	4.7
19	*****							12.6	12.1	11.7	11.2	10.3	7.9	4.6
20	*****								11.8	11.4	11.0	10.0	7.7	4.5
21	*****								11.5	11.1	10.7	9.8	7.6	4.4
22	*****								11.3	10.9	10.4	9.5	7.4	4.3
23	*****								11.0	10.6	10.2	9.3	7.2	4.2
24	*****									10.4	10.0	9.1	7.1	4.1
25	*****									10.2	9.8	8.9	6.9	4.0
30	*****										8.9	8.2	6.3	3.7
35	*****											7.6	5.9	3.4
40	*****												5.5	3.2
45	*****												5.2	3.0
50	*****												4.9	2.8
55	*****												4.7	2.7
60	*****													2.6
65	*****													2.5
70	*****													2.4

NOTE: FOR CORRECT USAGE OF THESE TABLES PLEASE REFER TO MICRODATA DOCUMENTATION

2000 Survey of Self-employment – User Guide

2000 Survey of Self-employment

Approximate Sampling Variability Tables for Saskatchewan

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	*****	59.3	58.4	56.8	55.2	53.5	51.8	50.1	48.3	46.4	42.3	32.8	18.9	
2	*****	41.3	40.2	39.0	37.9	36.7	35.4	34.1	32.8	29.9	23.2	13.4		
3	*****	33.7	32.8	31.9	30.9	29.9	28.9	27.9	26.8	24.4	18.9	10.9		
4	*****	29.2	28.4	27.6	26.8	25.9	25.0	24.1	23.2	21.2	16.4	9.5		
5	*****	25.4	24.7	23.9	23.2	22.4	21.6	20.7	18.9	14.7	8.5			
6	*****	23.2	22.5	21.9	21.2	20.4	19.7	18.9	17.3	13.4	7.7			
7	*****	21.5	20.9	20.2	19.6	18.9	18.2	17.5	16.0	12.4	7.2			
8	*****	20.1	19.5	18.9	18.3	17.7	17.1	16.4	15.0	11.6	6.7			
9	*****	18.9	18.4	17.8	17.3	16.7	16.1	15.5	14.1	10.9	6.3			
10	*****	17.5	16.9	16.4	15.8	15.3	14.7	13.4	10.4	6.0				
11	*****	16.6	16.1	15.6	15.1	14.6	14.0	12.8	9.9	5.7				
12	*****	15.9	15.5	15.0	14.5	13.9	13.4	12.2	9.5	5.5				
13	*****	15.3	14.9	14.4	13.9	13.4	12.9	11.7	9.1	5.3				
14	*****	14.3	13.9	13.4	12.9	12.5	12.4	11.3	8.8	5.1				
15	*****	13.8	13.4	12.9	12.5	12.0	10.9	8.5	4.9					
16	*****	13.4	13.0	12.5	12.1	11.6	10.6	8.2	4.7					
17	*****	13.0	12.6	12.1	11.7	11.2	10.3	8.0	4.6					
18	*****	12.6	12.2	11.8	11.4	10.9	10.0	7.7	4.5					
19	*****	11.9	11.5	11.1	10.6	9.7	7.5	4.3						
20	*****	11.6	11.2	10.8	10.4	9.5	7.3	4.2						
21	*****	11.3	10.9	10.5	10.1	9.2	7.2	4.1						
22	*****	11.1	10.7	10.3	9.9	9.0	7.0	4.0						
23	*****	10.4	10.1	9.7	8.8	6.8	3.9							
24	*****	10.2	9.9	9.5	8.6	6.7	3.9							
25	*****	10.0	9.7	9.3	8.5	6.6	3.8							
30	*****	8.8	8.5	7.7	6.0	3.5								
35	*****	7.8	7.2	5.5	3.2									
40	*****	6.7	5.2	3.0										
45	*****	6.3	4.9	2.8										
50	*****	4.6	2.7											
55	*****	4.4	2.6											
60	*****	4.2	2.4											
65	*****	2.3												
70	*****	2.3												
75	*****	2.2												
80	*****	2.1												

NOTE: FOR CORRECT USAGE OF THESE TABLES PLEASE REFER TO MICRODATA DOCUMENTATION

2000 Survey of Self-employment – User Guide

2000 Survey of Self-employment

Approximate Sampling Variability Tables for Alberta

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	*****	95.1	94.6	93.2	90.7	88.1	85.5	82.8	80.0	77.1	74.0	67.6	52.4	30.2
2	*****	67.3	66.9	65.9	64.1	62.3	60.5	58.5	56.6	54.5	52.4	47.8	37.0	21.4
3	*****	*****	54.6	53.8	52.4	50.9	49.4	47.8	46.2	44.5	42.7	39.0	30.2	17.5
4	*****	*****	47.3	46.6	45.3	44.1	42.7	41.4	40.0	38.5	37.0	33.8	26.2	15.1
5	*****	*****	*****	41.7	40.6	39.4	38.2	37.0	35.8	34.5	33.1	30.2	23.4	13.5
6	*****	*****	*****	38.0	37.0	36.0	34.9	33.8	32.6	31.5	30.2	27.6	21.4	12.3
7	*****	*****	*****	35.2	34.3	33.3	32.3	31.3	30.2	29.1	28.0	25.5	19.8	11.4
8	*****	*****	*****	32.9	32.1	31.2	30.2	29.3	28.3	27.2	26.2	23.9	18.5	10.7
9	*****	*****	*****	31.1	30.2	29.4	28.5	27.6	26.7	25.7	24.7	22.5	17.5	10.1
10	*****	*****	*****	29.5	28.7	27.9	27.0	26.2	25.3	24.4	23.4	21.4	16.6	9.6
11	*****	*****	*****	28.1	27.3	26.6	25.8	25.0	24.1	23.2	22.3	20.4	15.8	9.1
12	*****	*****	*****	26.9	26.2	25.4	24.7	23.9	23.1	22.2	21.4	19.5	15.1	8.7
13	*****	*****	*****	*****	25.2	24.4	23.7	23.0	22.2	21.4	20.5	18.7	14.5	8.4
14	*****	*****	*****	*****	24.2	23.6	22.8	22.1	21.4	20.6	19.8	18.1	14.0	8.1
15	*****	*****	*****	*****	23.4	22.8	22.1	21.4	20.6	19.9	19.1	17.5	13.5	7.8
16	*****	*****	*****	*****	22.7	22.0	21.4	20.7	20.0	19.3	18.5	16.9	13.1	7.6
17	*****	*****	*****	*****	22.0	21.4	20.7	20.1	19.4	18.7	18.0	16.4	12.7	7.3
18	*****	*****	*****	*****	21.4	20.8	20.2	19.5	18.9	18.2	17.5	15.9	12.3	7.1
19	*****	*****	*****	*****	20.8	20.2	19.6	19.0	18.3	17.7	17.0	15.5	12.0	6.9
20	*****	*****	*****	*****	20.3	19.7	19.1	18.5	17.9	17.2	16.6	15.1	11.7	6.8
21	*****	*****	*****	*****	19.8	19.2	18.7	18.1	17.5	16.8	16.2	14.7	11.4	6.6
22	*****	*****	*****	*****	19.3	18.8	18.2	17.6	17.1	16.4	15.8	14.4	11.2	6.4
23	*****	*****	*****	*****	18.9	18.4	17.8	17.3	16.7	16.1	15.4	14.1	10.9	6.3
24	*****	*****	*****	*****	18.5	18.0	17.5	16.9	16.3	15.7	15.1	13.8	10.7	6.2
25	*****	*****	*****	*****	*****	17.6	17.1	16.6	16.0	15.4	14.8	13.5	10.5	6.0
30	*****	*****	*****	*****	*****	16.1	15.6	15.1	14.6	14.1	13.5	12.3	9.6	5.5
35	*****	*****	*****	*****	*****	14.9	14.5	14.0	13.5	13.0	12.5	11.4	8.8	5.1
40	*****	*****	*****	*****	*****	*****	13.5	13.1	12.6	12.2	11.7	10.7	8.3	4.8
45	*****	*****	*****	*****	*****	*****	12.7	12.3	11.9	11.5	11.0	10.1	7.8	4.5
50	*****	*****	*****	*****	*****	*****	*****	11.7	11.3	10.9	10.5	9.6	7.4	4.3
55	*****	*****	*****	*****	*****	*****	*****	11.2	10.8	10.4	10.0	9.1	7.1	4.1
60	*****	*****	*****	*****	*****	*****	*****	10.7	10.3	9.9	9.6	8.7	6.8	3.9
65	*****	*****	*****	*****	*****	*****	*****	9.9	9.6	9.2	8.8	8.1	6.3	3.6
70	*****	*****	*****	*****	*****	*****	*****	9.6	9.2	8.8	8.4	7.8	6.0	3.5
75	*****	*****	*****	*****	*****	*****	*****	*****	8.9	8.5	8.1	7.6	5.9	3.4
80	*****	*****	*****	*****	*****	*****	*****	*****	8.6	8.3	7.9	7.4	5.7	3.3
85	*****	*****	*****	*****	*****	*****	*****	*****	8.4	8.0	7.6	7.1	5.5	3.2
90	*****	*****	*****	*****	*****	*****	*****	*****	*****	7.8	7.4	6.9	5.4	3.1
95	*****	*****	*****	*****	*****	*****	*****	*****	*****	7.6	7.2	6.7	5.2	3.0
100	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	6.8	6.4	4.9	2.7
125	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	4.7	2.5
150	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	4.3	2.1
200	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	2.1

NOTE: FOR CORRECT USAGE OF THESE TABLES PLEASE REFER TO MICRODATA DOCUMENTATION

2000 Survey of Self-employment – User Guide

2000 Survey of Self-employment

Approximate Sampling Variability Tables for Prairie provinces

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	*****	81.8	81.4	80.1	78.0	75.8	73.5	71.2	68.8	66.3	63.7	58.1	45.0	26.0
2	*****	57.9	57.6	56.7	55.2	53.6	52.0	50.4	48.6	46.9	45.0	41.1	31.8	18.4
3	*****	47.2	47.0	46.3	45.0	43.8	42.5	41.1	39.7	38.3	36.8	33.6	26.0	15.0
4	*****	40.9	40.7	40.1	39.0	37.9	36.8	35.6	34.4	33.1	31.8	29.1	22.5	13.0
5	*****	*****	36.4	35.8	34.9	33.9	32.9	31.8	30.8	29.6	28.5	26.0	20.1	11.6
6	*****	*****	33.2	32.7	31.8	30.9	30.0	29.1	28.1	27.1	26.0	23.7	18.4	10.6
7	*****	*****	30.8	30.3	29.5	28.7	27.8	26.9	26.0	25.1	24.1	22.0	17.0	9.8
8	*****	*****	28.8	28.3	27.6	26.8	26.0	25.2	24.3	23.4	22.5	20.6	15.9	9.2
9	*****	*****	*****	26.7	26.0	25.3	24.5	23.7	22.9	22.1	21.2	19.4	15.0	8.7
10	*****	*****	*****	25.3	24.7	24.0	23.3	22.5	21.8	21.0	20.1	18.4	14.2	8.2
11	*****	*****	*****	24.2	23.5	22.9	22.2	21.5	20.7	20.0	19.2	17.5	13.6	7.8
12	*****	*****	*****	23.1	22.5	21.9	21.2	20.6	19.9	19.1	18.4	16.8	13.0	7.5
13	*****	*****	*****	22.2	21.6	21.0	20.4	19.8	19.1	18.4	17.7	16.1	12.5	7.2
14	*****	*****	*****	21.4	20.8	20.3	19.7	19.0	18.4	17.7	17.0	15.5	12.0	6.9
15	*****	*****	*****	20.7	20.1	19.6	19.0	18.4	17.8	17.1	16.4	15.0	11.6	6.7
16	*****	*****	*****	20.0	19.5	19.0	18.4	17.8	17.2	16.6	15.9	14.5	11.3	6.5
17	*****	*****	*****	19.4	18.9	18.4	17.8	17.3	16.7	16.1	15.4	14.1	10.9	6.3
18	*****	*****	*****	18.9	18.4	17.9	17.3	16.8	16.2	15.6	15.0	13.7	10.6	6.1
19	*****	*****	*****	18.4	17.9	17.4	16.9	16.3	15.8	15.2	14.6	13.3	10.3	6.0
20	*****	*****	*****	17.9	17.4	17.0	16.4	15.9	15.4	14.8	14.2	13.0	10.1	5.8
21	*****	*****	*****	*****	17.0	16.5	16.0	15.5	15.0	14.5	13.9	12.7	9.8	5.7
22	*****	*****	*****	*****	16.6	16.2	15.7	15.2	14.7	14.1	13.6	12.4	9.6	5.5
23	*****	*****	*****	*****	16.3	15.8	15.3	14.8	14.3	13.8	13.3	12.1	9.4	5.4
24	*****	*****	*****	*****	15.9	15.5	15.0	14.5	14.0	13.5	13.0	11.9	9.2	5.3
25	*****	*****	*****	*****	15.6	15.2	14.7	14.2	13.8	13.3	12.7	11.6	9.0	5.2
30	*****	*****	*****	*****	14.2	13.8	13.4	13.0	12.6	12.1	11.6	10.6	8.2	4.7
35	*****	*****	*****	*****	13.2	12.8	12.4	12.0	11.6	11.2	10.8	9.8	7.6	4.4
40	*****	*****	*****	*****	12.3	12.0	11.6	11.3	10.9	10.5	10.1	9.2	7.1	4.1
45	*****	*****	*****	*****	*****	11.3	11.0	10.6	10.3	9.9	9.5	8.7	6.7	3.9
50	*****	*****	*****	*****	*****	10.7	10.4	10.1	9.7	9.4	9.0	8.2	6.4	3.7
55	*****	*****	*****	*****	*****	10.2	9.9	9.6	9.3	8.9	8.6	7.8	6.1	3.5
60	*****	*****	*****	*****	*****	9.8	9.5	9.2	8.9	8.6	8.2	7.5	5.8	3.4
65	*****	*****	*****	*****	*****	*****	9.1	8.8	8.5	8.2	7.9	7.2	5.6	3.2
70	*****	*****	*****	*****	*****	*****	8.8	8.5	8.2	7.9	7.6	6.9	5.4	3.1
75	*****	*****	*****	*****	*****	*****	8.5	8.2	7.9	7.7	7.4	6.7	5.2	3.0
80	*****	*****	*****	*****	*****	*****	8.2	8.0	7.7	7.4	7.1	6.5	5.0	2.9
85	*****	*****	*****	*****	*****	*****	7.7	7.5	7.2	7.0	6.7	6.1	4.7	2.7
90	*****	*****	*****	*****	*****	*****	7.5	7.3	7.0	6.7	6.4	5.8	4.4	2.6
95	*****	*****	*****	*****	*****	*****	7.3	7.1	6.8	6.5	6.2	5.6	4.2	2.5
100	*****	*****	*****	*****	*****	*****	7.1	6.9	6.6	6.4	6.1	5.5	4.1	2.4
125	*****	*****	*****	*****	*****	*****	*****	6.2	5.9	5.7	5.4	4.8	3.4	2.3
150	*****	*****	*****	*****	*****	*****	*****	*****	5.2	4.9	4.7	4.1	2.8	2.1
200	*****	*****	*****	*****	*****	*****	*****	*****	*****	4.1	3.8	3.2	2.2	1.8
250	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	2.8	2.2	1.6	1.3
300	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	1.5	1.1	0.9
350	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	1.0	0.8

NOTE: FOR CORRECT USAGE OF THESE TABLES PLEASE REFER TO MICRODATA DOCUMENTATION

2000 Survey of Self-employment – User Guide

2000 Survey of Self-employment

Approximate Sampling Variability Tables for British Columbia

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	*****	100.9	100.4	98.9	96.3	93.5	90.7	87.9	84.9	81.8	78.6	71.7	55.6	32.1
2	*****	71.4	71.0	69.9	68.1	66.1	64.2	62.1	60.0	57.8	55.6	50.7	39.3	22.7
3	*****	58.3	58.0	57.1	55.6	54.0	52.4	50.7	49.0	47.2	45.4	41.4	32.1	18.5
4	*****	*****	50.2	49.4	48.1	46.8	45.4	43.9	42.4	40.9	39.3	35.9	27.8	16.0
5	*****	*****	44.9	44.2	43.0	41.8	40.6	39.3	38.0	36.6	35.1	32.1	24.9	14.3
6	*****	*****	41.0	40.4	39.3	38.2	37.0	35.9	34.7	33.4	32.1	29.3	22.7	13.1
7	*****	*****	*****	37.4	36.4	35.4	34.3	33.2	32.1	30.9	29.7	27.1	21.0	12.1
8	*****	*****	*****	35.0	34.0	33.1	32.1	31.1	30.0	28.9	27.8	25.4	19.6	11.3
9	*****	*****	*****	33.0	32.1	31.2	30.2	29.3	28.3	27.3	26.2	23.9	18.5	10.7
10	*****	*****	*****	31.3	30.4	29.6	28.7	27.8	26.8	25.9	24.9	22.7	17.6	10.1
11	*****	*****	*****	29.8	29.0	28.2	27.4	26.5	25.6	24.7	23.7	21.6	16.8	9.7
12	*****	*****	*****	28.5	27.8	27.0	26.2	25.4	24.5	23.6	22.7	20.7	16.0	9.3
13	*****	*****	*****	27.4	26.7	25.9	25.2	24.4	23.5	22.7	21.8	19.9	15.4	8.9
14	*****	*****	*****	26.4	25.7	25.0	24.3	23.5	22.7	21.9	21.0	19.2	14.9	8.6
15	*****	*****	*****	25.5	24.9	24.2	23.4	22.7	21.9	21.1	20.3	18.5	14.3	8.3
16	*****	*****	*****	24.1	23.4	22.7	22.0	21.2	20.4	19.6	18.8	17.9	13.9	8.0
17	*****	*****	*****	23.3	22.7	22.0	21.3	20.6	19.8	19.1	18.3	17.4	13.5	7.8
18	*****	*****	*****	22.7	22.0	21.4	20.7	20.0	19.3	18.5	17.7	16.9	13.1	7.6
19	*****	*****	*****	22.1	21.5	20.8	20.2	19.5	18.8	18.0	17.2	16.5	12.7	7.4
20	*****	*****	*****	21.5	20.9	20.3	19.6	19.0	18.3	17.6	16.9	16.0	12.4	7.2
21	*****	*****	*****	21.0	20.4	19.8	19.2	18.5	17.8	17.1	16.4	15.7	12.1	7.0
22	*****	*****	*****	20.5	19.9	19.3	18.7	18.1	17.4	16.8	16.1	15.3	11.8	6.8
23	*****	*****	*****	20.1	19.5	18.9	18.3	17.7	17.1	16.4	15.8	15.0	11.6	6.7
24	*****	*****	*****	19.6	19.1	18.5	17.9	17.3	16.7	16.0	15.4	14.6	11.3	6.5
25	*****	*****	*****	19.3	18.7	18.1	17.6	17.0	16.4	15.7	15.1	14.3	11.1	6.4
30	*****	*****	*****	17.6	17.1	16.6	16.0	15.5	14.9	14.3	13.7	13.1	10.1	5.9
35	*****	*****	*****	*****	15.8	15.3	14.9	14.3	13.8	13.3	12.7	12.1	9.4	5.4
40	*****	*****	*****	*****	14.8	14.3	13.9	13.4	12.9	12.4	11.9	11.3	8.8	5.1
45	*****	*****	*****	*****	13.9	13.5	13.1	12.7	12.2	11.7	11.2	10.7	8.3	4.8
50	*****	*****	*****	*****	*****	12.8	12.4	12.0	11.6	11.1	10.6	10.1	7.9	4.5
55	*****	*****	*****	*****	*****	12.2	11.8	11.4	11.0	10.6	10.1	9.7	7.5	4.3
60	*****	*****	*****	*****	*****	11.7	11.3	11.0	10.6	10.1	9.7	9.3	7.2	4.1
65	*****	*****	*****	*****	*****	10.9	10.5	10.1	9.7	9.3	8.9	8.5	6.9	4.0
70	*****	*****	*****	*****	*****	10.5	10.1	9.7	9.3	8.9	8.5	8.1	6.6	3.8
75	*****	*****	*****	*****	*****	10.1	9.8	9.4	9.1	8.7	8.3	7.9	6.4	3.7
80	*****	*****	*****	*****	*****	9.5	9.1	8.8	8.4	8.0	7.6	7.2	6.0	3.6
85	*****	*****	*****	*****	*****	9.2	8.9	8.5	8.1	7.7	7.3	6.9	5.8	3.5
90	*****	*****	*****	*****	*****	8.9	8.6	8.2	7.8	7.4	7.0	6.6	5.5	3.4
95	*****	*****	*****	*****	*****	8.4	8.1	7.7	7.3	6.9	6.5	6.1	5.0	3.3
100	*****	*****	*****	*****	*****	8.2	7.9	7.5	7.1	6.7	6.3	5.9	4.9	3.2
125	*****	*****	*****	*****	*****	*****	7.0	6.7	6.3	5.9	5.5	5.1	4.1	2.9
150	*****	*****	*****	*****	*****	*****	*****	5.9	5.5	5.1	4.7	4.3	3.4	2.6
200	*****	*****	*****	*****	*****	*****	*****	*****	4.9	4.5	4.1	3.7	2.8	2.3
250	*****	*****	*****	*****	*****	*****	*****	*****	*****	4.0	3.6	3.2	2.4	2.0

NOTE: FOR CORRECT USAGE OF THESE TABLES PLEASE REFER TO MICRODATA DOCUMENTATION

2000 Survey of Self-employment – User Guide

2000 Survey of Self-employment

Approximate Sampling Variability Tables for Canada

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	101.6	101.1	100.6	99.1	96.4	93.7	90.9	88.0	85.0	82.0	78.7	71.9	55.7	32.1
2	71.8	71.5	71.2	70.1	68.2	66.3	64.3	62.2	60.1	57.9	55.7	50.8	39.4	22.7
3	*****	58.4	58.1	57.2	55.7	54.1	52.5	50.8	49.1	47.3	45.5	41.5	32.1	18.6
4	*****	50.6	50.3	49.5	48.2	46.9	45.5	44.0	42.5	41.0	39.4	35.9	27.8	16.1
5	*****	45.2	45.0	44.3	43.1	41.9	40.7	39.4	38.0	36.6	35.2	32.1	24.9	14.4
6	*****	41.3	41.1	40.4	39.4	38.3	37.1	35.9	34.7	33.5	32.1	29.3	22.7	13.1
7	*****	38.2	38.0	37.4	36.4	35.4	34.4	33.3	32.1	31.0	29.8	27.2	21.0	12.1
8	*****	35.8	35.6	35.0	34.1	33.1	32.1	31.1	30.1	29.0	27.8	25.4	19.7	11.4
9	*****	33.7	33.5	33.0	32.1	31.2	30.3	29.3	28.3	27.3	26.2	24.0	18.6	10.7
10	*****	32.0	31.8	31.3	30.5	29.6	28.8	27.8	26.9	25.9	24.9	22.7	17.6	10.2
11	*****	30.5	30.3	29.9	29.1	28.3	27.4	26.5	25.6	24.7	23.7	21.7	16.8	9.7
12	*****	29.2	29.0	28.6	27.8	27.1	26.2	25.4	24.6	23.7	22.7	20.7	16.1	9.3
13	*****	28.1	27.9	27.5	26.7	26.0	25.2	24.4	23.6	22.7	21.8	19.9	15.4	8.9
14	*****	27.0	26.9	26.5	25.8	25.0	24.3	23.5	22.7	21.9	21.0	19.2	14.9	8.6
15	*****	26.1	26.0	25.6	24.9	24.2	23.5	22.7	22.0	21.2	20.3	18.6	14.4	8.3
16	*****	25.3	25.2	24.8	24.1	23.4	22.7	22.0	21.3	20.5	19.7	18.0	13.9	8.0
17	*****	24.5	24.4	24.0	23.4	22.7	22.1	21.4	20.6	19.9	19.1	17.4	13.5	7.8
18	*****	23.8	23.7	23.4	22.7	22.1	21.4	20.7	20.0	19.3	18.6	16.9	13.1	7.6
19	*****	23.2	23.1	22.7	22.1	21.5	20.9	20.2	19.5	18.8	18.1	16.5	12.8	7.4
20	*****	22.6	22.5	22.2	21.6	21.0	20.3	19.7	19.0	18.3	17.6	16.1	12.4	7.2
21	*****	*****	22.0	21.6	21.0	20.5	19.8	19.2	18.6	17.9	17.2	15.7	12.1	7.0
22	*****	*****	21.5	21.1	20.6	20.0	19.4	18.8	18.1	17.5	16.8	15.3	11.9	6.9
23	*****	*****	21.0	20.7	20.1	19.5	19.0	18.4	17.7	17.1	16.4	15.0	11.6	6.7
24	*****	*****	20.5	20.2	19.7	19.1	18.6	18.0	17.4	16.7	16.1	14.7	11.4	6.6
25	*****	*****	20.1	19.8	19.3	18.7	18.2	17.6	17.0	16.4	15.7	14.4	11.1	6.4
30	*****	*****	18.4	18.1	17.6	17.1	16.6	16.1	15.5	15.0	14.4	13.1	10.2	5.9
35	*****	*****	17.0	16.7	16.3	15.8	15.4	14.9	14.4	13.9	13.3	12.1	9.4	5.4
40	*****	*****	15.9	15.7	15.2	14.8	14.4	13.9	13.4	13.0	12.4	11.4	8.8	5.1
45	*****	*****	14.8	14.8	14.4	14.0	13.6	13.1	12.7	12.2	11.7	10.7	8.3	4.8
50	*****	*****	14.0	14.0	13.6	13.3	12.9	12.4	12.0	11.6	11.1	10.2	7.9	4.5
55	*****	*****	13.4	13.4	13.0	12.6	12.3	11.9	11.5	11.1	10.6	9.7	7.5	4.3
60	*****	*****	12.8	12.8	12.4	12.1	11.7	11.4	11.0	10.6	10.2	9.3	7.2	4.1
65	*****	*****	12.3	12.3	12.0	11.6	11.3	10.9	10.5	10.2	9.8	8.9	6.9	4.0
70	*****	*****	11.8	11.8	11.5	11.2	10.9	10.5	10.2	9.8	9.4	8.6	6.7	3.8
75	*****	*****	11.4	11.4	11.1	10.8	10.5	10.2	9.8	9.5	9.1	8.3	6.4	3.7
80	*****	*****	11.1	11.1	10.8	10.5	10.2	9.8	9.5	9.2	8.8	8.0	6.2	3.6
85	*****	*****	10.7	10.7	10.5	10.2	9.9	9.5	9.2	8.9	8.5	7.8	6.0	3.5
90	*****	*****	10.4	10.4	10.2	9.9	9.6	9.3	9.0	8.6	8.3	7.6	5.9	3.4
95	*****	*****	10.2	10.2	9.9	9.6	9.3	9.0	8.7	8.4	8.1	7.4	5.7	3.3
100	*****	*****	9.9	9.9	9.6	9.4	9.1	8.8	8.5	8.2	7.9	7.2	5.6	3.2
125	*****	*****	*****	8.6	8.4	8.1	7.9	7.6	7.3	7.0	6.7	6.4	5.0	2.9
150	*****	*****	*****	7.9	7.7	7.4	7.2	6.9	6.7	6.4	6.1	5.9	4.5	2.6
200	*****	*****	*****	6.8	6.6	6.4	6.2	6.0	5.8	5.6	5.4	5.1	3.9	2.3
250	*****	*****	*****	*****	5.9	5.8	5.6	5.4	5.2	5.0	4.8	4.5	3.5	2.0
300	*****	*****	*****	*****	5.4	5.2	5.1	4.9	4.7	4.5	4.4	4.1	3.2	1.9
350	*****	*****	*****	*****	*****	4.9	4.7	4.5	4.4	4.2	4.1	3.8	3.0	1.7
400	*****	*****	*****	*****	*****	4.5	4.4	4.3	4.1	3.9	3.7	3.4	2.8	1.6
450	*****	*****	*****	*****	*****	*****	4.1	4.0	3.9	3.7	3.5	3.2	2.6	1.5
500	*****	*****	*****	*****	*****	*****	3.9	3.8	3.7	3.5	3.4	3.2	2.5	1.4
750	*****	*****	*****	*****	*****	*****	*****	2.9	2.8	2.6	2.5	2.3	1.8	1.0
1000	*****	*****	*****	*****	*****	*****	*****	*****	2.3	2.1	2.0	1.8	1.4	0.8
1500	*****	*****	*****	*****	*****	*****	*****	*****	*****	1.8	1.7	1.5	1.1	0.6

NOTE: FOR CORRECT USAGE OF THESE TABLES PLEASE REFER TO MICRODATA DOCUMENTATION

11.0 Weighting

Since the Survey of Self-employment used a sub-sample of the LFS sample, the derivation of weights for the survey records is clearly tied to the weighting procedure used for the LFS. The LFS weighting procedure is briefly described below.

11.1 Weighting Procedures for the LFS

In the LFS, the final weight attached to each record is the product of the following factors: the basic weight, the cluster sub-weight, the balancing factor for non-response, and the province-age-sex ratio adjustment factor. Each is described below.

Basic Weight

In a probability sample, the sample design itself determines weights which must be used to produce unbiased estimates of the population. Each record must be weighted by the inverse of the probability of selecting the person to whom the record refers. In the example of a 2% simple random sample, this probability would be 0.02 for each person and the records must be weighted by $1/0.02=50$. Because all eligible individuals in a dwelling are interviewed (directly or by proxy), this probability is essentially the same as the probability with which the dwelling is selected.

Cluster Sub-weight

The cluster delineation is such that the number of dwellings in the sample increases very slightly with moderate growth in the housing stock. Substantial growth can be tolerated in an isolated cluster before the additional sample represents a field collection problem. However, if growth takes place in more than one cluster in an interviewer assignment, the cumulative effect of all increases may create a workload problem. In clusters where substantial growth has taken place, sub-sampling is used as a means of keeping interviewer assignments manageable. The cluster sub-weight represents the inverse of this sub-sampling ratio in clusters where sub-sampling has occurred.

Non-response

Notwithstanding the strict controls of the LFS, some non-response is inevitable, despite all the attempts made by the interviewers. The LFS non-response rate was approximately 7% in April 2000. For certain types of non-response (eg. household temporarily absent, refusal), data from a previous month's interview with the household if any, is brought forward and used as the current month's data for the household.

In other cases, non-response is compensated for by proportionally increasing the weights of responding households. The weight of each responding record is increased by the ratio of the number of households that should have been interviewed, divided by the number that were actually interviewed. This adjustment is done separately for non-response areas, which are defined by employment insurance region, type of area, and rotation group. It is based on the assumption that the households that have been interviewed represent the characteristics of those that should have been interviewed. To the extent that this assumption is not true, the estimates will be somewhat biased.

LFS Sub-Weight

The product of the previously described weighting factors is called the LFS sub-weight. All members of the same sampled dwelling have the same sub-weight.

Subprovincial and Province-Age-Sex Adjustments

The sub-weight can be used to derive a valid estimate of any characteristic for which information is collected by the LFS. In particular, estimates are produced of the total number of persons 15+ in provincial economic regions and the 24 large metropolitan areas as well as of designated age-sex groups in each of the ten provinces.

Independent estimates are available monthly for various age and sex groups by province. These are population projections based on the most recent Census data, records of births and deaths, and estimates of migration. In the final step, this auxiliary information is used to transform the sub-weight into the final weight. This is done using a calibration method. This method ensures that the final weights it produces sum to the census projections for the auxiliary variables, namely various age-sex groups, economic regions, census metropolitan areas, and rotation groups.

This weighting procedure ensures consistency with external Census counts and that each rotation group is representative of the population, and also ensures that every member of the economic family is assigned the same weight.

11.2 Weighting Procedures for the Survey of Self-employment

The principles behind the calculation of the weights for the SSE are identical to those for the LFS. However, further adjustments are made to the LFS weights in order to derive a final weight for the individual respondent records on the SSE microdata file.

- (1) An adjustment to account for the use of only five of the six LFS rotation groups for the SSE
- (2) An adjustment to account for the additional non-response to the supplementary survey i.e., non-response to the SSE for individuals who did respond to the LFS and were eligible to receive the SSE
- (3) An adjustment to calibrate the weights to province-age-sex totals for the self-employed population, after the above adjustments are made.

Adjustment (1) is done by multiplying the LFS sub-weight for all records eligible to receive the SSE by a factor of 6/5. Call this weight the SSE sub-sample weight.

Adjustment (2) is accounted for by multiplying the SSE sub-sample weight for all SSE respondents by a factor of

$$\frac{\textit{weighted total of SSE respondents and non-respondents}}{\textit{weighted total of SSE respondents}}$$

In this case the weighted totals are generated by using the SSE sub-sample weight. The resulting product of the SSE sub-sample weight and the adjustment is called the SSE sub-weight .

More than one factor was calculated using the formula described above. Within each province the dataset was divided up into two or more non-response groupings. These splits were based upon variables available from the LFS for both SSE respondents and non-respondents that appeared to be good predictors of high and low SSE non-response. The decision on the best variables to use was done through the use of a software called Knowledge Seeker. The groupings differed from one province to next. In total there were fifty-one non-response groupings across the ten provinces. An individual's sub-sample weight was then adjusted by the factor for the grouping that corresponded to the individual's characteristics.

Adjustment (3) is calculated by

$$\frac{\text{LFS self-employed population estimate for province-sex-age group } i}{\text{weighted total of SSE respondents for province-sex-age group } i}$$

where the weighted total is computed by using the SSE sub-weight. The SSE sub-weight is then multiplied by the corresponding adjustment factor to create a new weight called the SSE final weight. This is the weight that appears on the accompanying file. The total of the SSE final weights for each of the province-sex-age groups will match the corresponding totals from the LFS for the self-employed population.

In all provinces except those in the Atlantic region, there were eight sex-age groups - males and females split into age groups 15-34, 35-44, 45-54, 55+. In the Atlantic provinces, the two oldest age groups were combined into one (45+) for both males and females, resulting in only six groupings.

12.0 Questionnaires

12.1 The Labour Force Survey Questionnaire

The Household Demographics Questionnaire is used to list all household members whose usual place of residence is the selected dwelling. It is both a survey operations control document and a record of socio-demographic information on household members.

The Labour Force Survey Questionnaire (LFS_QuestE.pdf) is used to list all household members whose usual place of residence is the selected dwelling. It is also used to collect information on the current and most recent labour market activity of all household members 15 years of age or older. It includes questions on hours of work, job tenure, type of work, reason for hours lost or absent, job search undertaken, availability for work, and school attendance.

12.2 The Survey of Self-Employment Questionnaires

The Self-employment questionnaire was used in April 2000 to collect the information for the supplementary survey. The file SSE2000_QuestE.pdf contains the English questionnaire.

13.0 File Layout with Univariate Frequencies

The users should be aware that the counts vary depending on the flow of questions. Whenever possible, where applicable, the skip patterns have been identified in the file layout. Still, the user should consult the questionnaire to check the flow of questions.