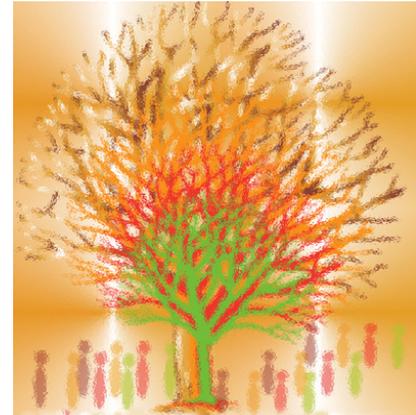




Catalogue no. 91-528-XIE

Population and Family Estimation Methods at Statistics Canada

Demography Division



 Statistics
Canada Statistique
Canada

Canada

How to obtain more information

Specific inquiries about this product and related statistics or services should be directed to: Development and Demographic-Methods Section, Demography Division, Statistics Canada, Ottawa, Ontario, K1A 0T6.

For information about this product or the wide range of services and data available from Statistics Canada, visit our website at www.statcan.ca or contact us by e-mail at infostats@statcan.ca or by telephone from 8:30 a.m. to 4:30 p.m. Monday to Friday:

Statistics Canada National Contact Centre

Toll-free telephone (Canada and the United States):

Inquiries line	1-800-263-1136
National telecommunications device for the hearing impaired	1-800-363-7629
Fax line	1-877-287-4369
Depository Services Program inquiries line	1-800-635-7943
Depository Services Program fax line	1-800-565-7757

Local or international calls:

Inquiries line	1-613-951-8116
Fax line	1-613-951-0581

Information to access the product

This product, Catalogue no. 91-528-XIE, is available for free in electronic format. To obtain a single issue, visit our website at www.statcan.ca and select "Publications."

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, the Agency has developed standards of service which its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll free at 1-800-263-1136. The service standards are also published on www.statcan.ca under "About us" > "Providing services to Canadians."



Statistics Canada
Demography Division

Population and Family Estimation Methods at Statistics Canada

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2007

All rights reserved. The content of this electronic publication may be reproduced, in whole or in part, and by any means, without further permission from Statistics Canada, subject to the following conditions: that it be done solely for the purposes of private study, research, criticism, review or newspaper summary, and/or for non-commercial purposes; and that Statistics Canada be fully acknowledged as follows: Source (or "Adapted from", if appropriate): Statistics Canada, year of publication, name of product, catalogue number, volume and issue numbers, reference period and page(s). Otherwise, no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form, by any means—electronic, mechanical or photocopy—or for any purposes without prior written permission of Licensing Services, Client Services Division, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6.

December 2007

Catalogue no. 91-528-XIE
ISBN: 978-0-662-47399-2

Frequency: Occasional

Ottawa

La version française de cette publication est disponible sur demande (n° 91-528-XIF au catalogue).

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued cooperation and goodwill.

Symbols

The following standard symbols are used in Statistics Canada publications:

- . not available for any reference period
- .. not available for a specific reference period
- ... not applicable
- 0 true zero or a value rounded to zero
- 0^s value rounded to 0 (zero) where there is a meaningful distinction between true zero and the value that was rounded
- ^p preliminary
- ^r revised
- x suppressed to meet the confidentiality requirements of the *Statistics Act*
- ^E use with caution
- F too unreliable to be published

Acknowledgements

This manual was prepared by the Development and Demographic Methods Section, Demography Division under the editorship of Shirley Loh, analyst in the section. The written work is based on numerous sources and represents the collaborative work of many individuals in Demography Division and other related divisions such as Social Survey Methods Division, Small Area and Administrative Data Division and Health Statistics Division at Statistics Canada. Demographic methods also rely on a number of administrative records administered by various government bodies responsible for managing these programs; namely the Canada Revenue Agency and Citizenship and Immigration Canada and the various provincial bodies responsible for reporting vital events to the bureau.

Inspired from previous editions of this report, the implicit contribution of many others is thus acknowledged. While many people provided valuable input or assistance, special recognition must be given to Johanne Denis (Assistant Director in Demography Division), André Cyr (Chief, Development and Demographic Methods Section) who spent countless hours editing and providing guidance and information in all stages of the development of the manual. Very important contributions were provided by Christian Thibault and Peter Dick (Social Survey Methods Division), Leslie Geran (Senior Analyst, Health Statistics Division), Denis Morissette, Julien Bérard-Chagnon and Peter Wilkinson (Development and Demographic Methods Section), Hélène Landry, Hubert Denis, Sylvie Bonhomme and Stéphane Gilbert (Demographic Estimates Section) who also served as a valuable resource for information pertaining to current estimation methods.

Finally, Julie Dussault (Development and Demographic Methods Section) provided technical expertise in the production of the document into its final format and, with the assistance of Janie Saumure provided a great deal of support to ensure that this product met the standards and guidelines for electronic publishing in Statistics Canada.

Table of contents

Page

Introduction

8

Chapter 1	Postcensal and intercensal population estimates, Canada, provinces and territories	13
Chapter 2	Base population (Censal estimates)	21
Chapter 3	Births and deaths	38
Chapter 4	Immigration	43
Chapter 5	Non-permanent residents	47
Chapter 6	Emigrants, net temporary emigrants and returning emigrants	53
Chapter 7	Interprovincial migration	65
Chapter 8	Subprovincial estimates of population	73
Chapter 9	Population estimates by age, sex and marital status	81
Chapter 10	Census family estimates	94

List of tables

Table I	Availability of estimates	10
Table 1.1	Sources and references of postcensal population estimates - Component data	15
Table 3.1	Births, deaths, natural increase, Canada, 1971/1972 to 2005/2006	39
Table 4.1	Annual number of immigrants, Canada, provinces and territories, 1971/1972 to 2005/2006	44
Table 9.1	Percent distribution of husband-wife families by age group of wife and age group of husband according to the legal marital status concept: Ontario, 2001 Census	89

Table of contents (concluded)

Page

List of tables (concluded)

Table 9.2	Calculation of the incidence of widowhood according to the legal marital status concept: An example for females aged 45 to 49 years, Ontario, 2001/2002	90
Table 10.1	Sources and references of data used to generate the postcensal estimates of census families	96
Table 10.2	Events and hypotheses used to reconstruct changes in the number of census families	97

List of figures

Figure 1.1	International migration flows for Canada	14
Figure 1.2	Method of distributing the error of closure by cohort	19
Figure 4.1	Immigrants as a percentage of total provincial and territorial population, 2001	45
Figure 5.1	NPR flows by broad type	48
Figure 8.1	Standard Geographical Classification (SGC) hierarchy	74
Figure 9.1	Transition from a distribution of demographic events by age and period to a distribution by age and birth cohort	83

Appendix I: Glossary	111
-----------------------------	-----

Appendix II: Symbols used in equations and their meaning	122
---	-----

References	125
-------------------	-----

Introduction

Population and family estimation methods at Statistics Canada

Demographic estimates are the cornerstone of statistical measurement for the population and are put to a wide variety of uses to better understand the economic and social situation in Canada. Their importance is underscored by the statutory requirements for the calculation of revenue transfers and cost-sharing programs between the various levels of government. As the National Statistical Agency responsible for these data, among others, timeliness, accuracy and reliability of these data are of the utmost importance.

Every five years, Statistics Canada conducts a national census of population, through which a wide range of demographic data is collected on the Canadian population. The census collects information on citizens, landed immigrants and non-permanent residents (NPR) residing in Canada. While some countries have a system of continuous population registration from which to obtain basic demographic data on the state and movement of the population for non-census years, Canada does not.

At a five year interval, the census cannot fill the need for timely data that are required by the various statistical programs in Statistics Canada as well as outside users of data. To fill this gap in the information system, Statistics Canada has a Population and Family Estimates Program. By using the most recent census data collected by Statistics Canada and administrative data provided by other government departments, various methodological techniques are used to produce estimates of the Canadian population between censuses. Moreover, Statistics Canada has also developed a Population Projections Program. Based on extrapolations of past trends, these projections reflect informed hypotheses of Canada's demographic future.

In 1987, Statistics Canada published its first manual detailing the procedures used for estimating population, called *Population Estimation Methods, Canada* (Catalogue no. 91-528). The manual represented a direct response to a long-standing policy at Statistics Canada that requires the methods used to produce the Bureau's statistical information to be open to public inspection. In this current version of the report, readers will notice a deviation from previously released versions as the quality indicators have been removed. Ongoing quality indicators are available in the annual and quarterly publications¹. The analysis of these quality indicators and special data quality evaluations will be featured on their own in a companion publication on data quality indicators and analysis.

In an effort to constantly improve the estimates of population, conceptual and methodological changes have been made to estimation procedures since the publication of the 1987 manual. In 1993, the decision to adjust the population counts for net census undercoverage and the addition of NPRs in the calculation of population estimates are two of the major methodology modifications. Other changes that necessitated updating the 1987 manual include the new approach to the calculation of preliminary postcensal estimates at subprovincial levels, and modifications to the calculation of certain components

1. Refer to **Quarterly Demographic Estimates**, catalogue no.91-002-X and **Annual Demographic Estimates**, catalogue no. 91-215-X.

of population change which are not readily available and require estimation, such as emigration. At times, changes in the methodology of the census or its concepts and definitions will necessitate a change in the demographic methods used for estimation.

This manual focuses on the methods used to produce population and family estimates. However, Demography Division also produces population projections. Population estimates and projections are similar in that they both indicate the size of the population for a given time. The distinction between them is related to their underlying concepts. Population estimates are approximate counts of the current (or recent) population, whereas population projections are expected population counts based on a defined set of assumptions. For information on population projection methods, please refer to *Population Projections for Canada, Provinces and Territories* (Catalogue 91-520), or *Population Projections of Visible Minority Groups, Canada, Provinces and Regions* (Catalogue 91-541), or *Projections of the Aboriginal Populations, Canada, Provinces and Territories* (Catalogue 91-547).

The literature on methodological changes in the calculation of population and family estimates can be found piecemeal throughout various Statistics Canada publications. This manual is intended to be a compendium of the methods and the current procedures used by Statistics Canada to produce and release population and family estimates.

Estimates

Essentially, there are two categories of estimates produced by Statistics Canada: postcensal and intercensal. The first category, postcensal estimates, are produced by using data from the most recent census (adjusted for net census undercoverage) and estimates of the components of demographic change since that last census. These components include births, deaths, immigration, total emigration, net change in non-permanent residents, and interprovincial migration. Another component, intraprovincial migration is relevant for estimates within subprovincial areas in Canada. All the elements of the production of population estimates will be discussed in this manual.

Three types of postcensal estimates are produced: preliminary, updated and final estimates, referring to the time frame in which they become available. Preliminary estimates are typically available within three to four months after the reference date while updated estimates are usually available within one year. Final estimates, the most accurate postcensal estimates available, typically take two years to complete. The productions of three types of estimates, while sometimes confusing to users, is the strategy that best satisfies the commitment of Statistics Canada to produce data in a timely and accurate manner.

Intercensal estimates are produced every five years and reconcile previous postcensal estimates with the latest census counts. They are generated as soon as census population counts become available. This process typically takes two years after census data collection to complete.

Disseminated level of detail for the estimates

Population estimates are disseminated at four geographic levels, including province and territory, census division, census metropolitan area and economic region. Demographic estimates at custom-defined subprovincial levels (e.g., census subdivision, urban centres, health regions, etc.) are possible and can be estimated through cost-recovery special tabulations. For timeliness, and because some

components of demographic growth are not available until several months after the reference date, three kinds of postcensal estimates are produced: preliminary postcensal (PP), updated postcensal (PR) and final postcensal (PD). Intercensal estimates are produced using postcensal estimates and counts from two consecutive censuses adjusted for net census undercoverage (NCU). According to the level of geography and type of estimates, different demographic characteristics of the population, including age, sex, or marital status are produced. Table I shows the level of detail and the reference period for which population estimates are disseminated.

Table I
Availability of estimates

Geography	Characteristics	Type of estimate*	Frequency
Canada, provinces and the territories	Total population	<ul style="list-style-type: none"> • Preliminary postcensal • Updated postcensal • Final postcensal • Intercensal 	Quarterly and annual estimates
	<ul style="list-style-type: none"> • Age • Sex • Marital status 	<ul style="list-style-type: none"> • Preliminary postcensal • Updated postcensal • Final postcensal • Intercensal 	Annual estimates
	Family <ul style="list-style-type: none"> • Family size and structure • Age of children • Age of husband and age of wife • Age and sex of lone parents • Persons in families 	<ul style="list-style-type: none"> • Preliminary postcensal • Updated postcensal • Final postcensal • Intercensal 	Annual
Census division / Census metropolitan area / Economic region	<ul style="list-style-type: none"> • Total population • Age • Sex 	<ul style="list-style-type: none"> • Preliminary postcensal • Updated postcensal • Final postcensal • Intercensal 	Annual
*All types of estimates are released by the various components of demographic growth except for intercensal estimates			

Demographic estimates at custom-defined subprovincial levels can be produced by applying synthetic estimation techniques, under the assumption that the larger geographic area's distribution still holds for lower or alternate geographies; or by regression estimation techniques, under the assumption that the known population totals, usually from the census, still hold.

Users of population and family estimates

In Statistics Canada, demographic information is used to calibrate sampling weights of many social and household based survey data, and is a fundamental part of the analytical framework for most statistical programs. Current estimates of Canada's population have other wide ranging applications in the areas of planning and program evaluation in both the public and private sectors. For example, the calculation of revenue transfers and grants under various federal statutory programs, as well as cost-sharing agreements between federal, provincial, territorial and municipal governments are highly dependent on demographic data. Statistics Canada is under statutory obligation to provide the federal government with annual population figures as well as various economic indicators (e.g., gross

domestic product) that have been certified by the Chief Statistician of Canada. These figures are used to determine the amounts payable under various federal-provincial fiscal arrangements, such as the Equalization Program and the Canada Health and Social Transfer (CHST), according to a per capita funding formula.

Equalization payments are important for the reduction of fiscal disparities among provinces. These payments, made by the federal government, enable less prosperous provincial governments to provide their residents with public services that are reasonably comparable to those in other provinces, at reasonably comparable levels of taxation. Payments are made to entitled² provinces in the form of a block grant, which allows the recipient provincial governments to spend the funds according to their own priorities.

The CHST (the largest federal transfer to provinces and territories) is also a block-funding program, allowing each of the provinces and territories to allocate spending among social programs such as health care, post-secondary education, social assistance and social services, according to their own priorities. In 2006–07 fiscal year, the Government of Canada provided support of \$61 billion to provincial and territorial governments through its major cash and tax transfers including the Canada Health Transfer (CHT), Canada Social Transfer, (CST), Wait Times Reduction Transfer, Equalization and Territorial Formula Financing³.

Population is a key variable in allocation formulae used by the federal government in defining its annual spending. As billions of dollars depend directly upon these allocation formulae, timely and accurate estimates are extremely important.

The Government of Canada relies upon population data to also inform some of its most fundamental policies. For example, up-to-date information on the evolving demographic situation in Canada serves useful in informing cabinet on its decisions relating to immigration policy. In the administration of public policy, current population figures also serve to increase the cost-effectiveness of program spending.

Provincial and municipal governments also use demographic estimates for planning social programs. These governments further use estimates to establish cost-sharing agreements on the basis of per capita spending formulae. Among the public service users of Statistics Canada's estimates, are education and public health planners, public administrators responsible for policing, criminal justice, municipal administration, and waste and environmental management.

Demographic estimates are fundamental in the calculation of social and economic indicators, including birth rates, death rates, school enrolment rates, unemployment rates, life expectancy, etc.

-
2. Equalization entitlements are calculated using a measure of “fiscal capacity” which compares each province’s ability to generate revenues at national average tax rates to those of a threshold or “standard”. The current standard (which has been used since 1982) uses the five “middle-income” provinces: Quebec, Ontario, Manitoba, Saskatchewan and British Columbia. Provinces with fiscal capacities below the standard are entitled to equalization payments to bring their capacity up to the standard (Department of Finance Canada, 2002). Budget 2007 introduced a new program legislated through 2013/14 to provide long-term predictability for provinces. The new program is based on the recommendations made by the Expert Panel on Equalization and Territorial Formula Financing (ref: O’Brien Report) (Department of Finance Canada, 2007; <http://www.fin.gc.ca/access/fedprove.html>).
 3. “Federal Transfers to Provinces and Territories”, on the Department of Finance Canada’s website (<http://www.fin.gc.ca>).

Population serves directly as the denominator in many of these indicators. Survey researchers, whether at Statistics Canada, academia, or in private polling agencies, must use up-to-date population figures in the planning of survey research and in the calculation of sampling weights.

The private sector also uses demographic estimates for business planning, marketing research and investment demographics. The estimates also help companies to properly segment and target its market, or make sound investment decisions.

Outline of this manual

The first chapter of this manual presents the overall methods used to produce population estimates and focuses on describing the various levels of estimation: postcensal and intercensal estimates. It introduces the concept of the base population and the factors of demographic growth and their components; from components that lead to “Natural increase”, that is births and deaths, to international effects like “International migration” (immigrants, emigrants, returning emigrants, net temporary emigration and net non-permanent residents), to redistributive effects such as “Interprovincial migration”.

The individual chapters that follow include discussions on data sources, relevant concepts and methodology for each of these components. Each chapter will describe how estimates are produced, what data are used in the calculations and where they come from. This should give the reader a clearer picture of the strengths and weaknesses of each procedure.

Chapter 2 describes how the base population is determined and adjusted following the most recent census of the population. Chapters 3 to 7 describe the components of population growth for Canada, the provinces and territories. Chapter 3 focuses on information on births and deaths. Chapters 4 through 6 present the components of international migration (immigration is presented in Chapter 4, non-permanent residents in Chapter 5, and Chapter 6 deals with the total emigration component). Interprovincial migration is discussed in Chapter 7.

Chapter 8 describes the methods used to produce subprovincial population estimates. Chapter 9 focuses on the population estimates by age, sex and marital status. Finally, Chapter 10 describes the methods for estimating census families.

A glossary of principal terms is included at the end of this manual (Appendix I), followed by notations used in formulae presented herein (Appendix II). An attempt was made to simplify and standardize the notation system in the formulae throughout this manual. As such, notations in the manual may differ from those in previous publications of population estimates.

Chapter 1

Postcensal and intercensal population estimates, Canada, provinces and territories

This chapter describes the types of method used by Statistics Canada to calculate postcensal and intercensal estimates for the total population and for the population by age and sex at the provincial and territorial levels. The sources of data used to produce these estimates are also given.

1.1 Postcensal population estimates, Canada, provinces and territories

1.1.1 Definition and calculation of provincial and territorial postcensal estimates of total population

Postcensal population estimates are produced using data from the most recent census (adjusted for net census undercoverage) and estimates of the components of population growth since that census. Following each census, the first postcensal estimate refers to the adjusted census population forwarded to July 1 by taking into account the components of population growth between Census Day and July 1 of the census year. The component method used to produce postcensal estimates is a population accounting system, where modifications are made to the current census population adjusted for net census undercoverage or most recent estimate by adding and subtracting the components of population growth that occur between July 1 and the reference date of the estimate.

The factors of demographic growth and their components are:

Natural increase

- births
- deaths

International migration

- immigrants
- emigrants
- returning emigrants
- net temporary emigration
- net non-permanent residents

Interprovincial migration

- in-migrants
- out-migrants

These components can also be divided into two groups, according to the type of data used: those components for which data are readily available, including births, deaths, and immigration, and those that have to be estimated, including interprovincial migration, emigrants, returning emigrants, net temporary emigration, and net non-permanent residents (NPRs).

The two components of natural increase, i.e. births and deaths, have similar methodological approach when it comes to estimation.

Growth due to international migration represent the movement of population between Canada and a foreign country which involves a change in the usual place of residence.

Figure 1.1
International migration flows for Canada



In the population estimation program, international migration consists of five components: immigration, emigration, returning emigrants, net temporary emigration and net non-permanent residents. International migration flows can be categorized as either permanent or temporary. Permanent flows are composed of persons arriving in Canada for permanent residence (landed immigrants), Canadian citizens or landed immigrants returning to Canada after previously emigrated from Canada (returning emigrants), and Canadian citizens or landed immigrants leaving Canada to establish a permanent residence in another country (emigrants). Temporary flows relate to foreigners arriving for temporary stay in Canada and leaving after their stay ends (non-permanent residents), as well as Canadian citizens and landed immigrants living temporarily abroad who have not maintained a usual place of residence in Canada (temporary emigration).

Net non-permanent residents represent the variation in the number of non-permanent residents between two dates, and net temporary emigration represents the variation in the number of temporary emigrants between two dates. Different methodological approaches are used; one for the immigration component, another one for non-permanent residents and a more model based approach for the remaining components of international change (emigration, returning emigrants, and net temporary emigration).

Finally, the last factor of demographic growth that is discussed is the interprovincial migration. While this factor does not affect the total Canadian population it does affect the regional population counts and is a significant issue for the estimation program.

Table 1.1 shows the sources and references of component data used to generate the postcensal population.

Table 1.1
Sources and references of postcensal population estimates - Component data

Component	Sources and references (if applicable)
Base Population (Censal estimate)	Census population counts (Census of Canada, Catalogue no. 93F0051XIE) Data from Reverse Record Check (RRC) (<i>Coverage, 2001 Census Technical Report (Reference Products 2001 Census)</i> , Catalogue no. 92-394-XIE, Statistics Canada)
Births and deaths	Health Statistics Division, Statistics Canada Demography Division, Statistics Canada (<i>Quarterly Demographic Estimates</i> , Catalogue no. 91-002-XWE)
Marriages and divorces	Health Statistics Division, Statistics Canada
Immigration	Citizenship and Immigration Canada (CIC) data
Emigration	Data from the Canada Child Tax Benefit program (CCTB) collected by the Canada Revenue Agency (CRA); data from the T1 Family File (T1FF) derived from the CRA's T1 file by Small Area and Administrative Data Division (SAADD) of Statistics Canada, and data from the Office of Immigration Statistics, U.S. Department of Homeland Security
Returning emigrants	Data from the CRA's CCTB file and the SAADD's T1FF
Net temporary emigration	Data from the RRC, 1996 and 2001 censuses of Canada (<i>Coverage, 1996 Census Technical Report</i> , Catalogue no. 92-370-XIE, and <i>Coverage, 2001 Census Technical Report (Reference Products 2001 Census)</i> , Catalogue no. 92-394-XIE, Statistics Canada)
Net non-permanent residents	CIC data
Interprovincial and subprovincial migration	Data from the CRA's CCTB file and the SAADD's T1FF

Estimates of population are first produced for each province and territory, and then summed to obtain an estimate of the population of Canada.

The component method used in estimating total provincial and territorial populations is expressed as follows:

Equation 1.1:

$$P_{(t+1)} = P_t + B_{(t,t+1)} - D_{(t,t+1)} + I_{(t,t+1)} - (E_{(t,t+1)} + \Delta TE_{(t,t+1)}) + RE_{(t,t+1)} + \Delta NPR_{(t,t+1)} + \Delta N_{(t,t+1)}$$

where for any given province and territory

- ($t, t + 1$) = interval between the date of the previous reference period, time t , and the reference date of the estimate, time $t+1$;
- $P_{(t+1)}$ = estimate of population at time $t + I$;
- P_t = base population at time t (census adjusted for net census undercoverage or most recent estimates);
- $B_{(t,t+1)}$ = number of births between time t and $t + I$;
- $D_{(t,t+1)}$ = number of deaths between time t and $t + I$;
- $I_{(t,t+1)}$ = number of immigrants between time t and $t + I$;
- $E_{(t,t+1)}$ = number of emigrants between time t and $t + I$;
- $\Delta TE_{(t,t+1)}$ = net temporary emigration between time t and $t + I$;
- $RE_{(t,t+1)}$ = number of returning emigrants between time t and $t + I$;
- $\Delta NPR_{(t,t+1)}$ = net non-permanent residents between time t and $t + I$ (can be either positive or negative);
- $\Delta N_{(t,t+1)}$ = net interprovincial migration between time t and $t + I$ (can be either positive or negative);

1.1.2 Provincial and territorial postcensal population estimates by age and sex

Postcensal estimates of the population by age and sex are produced using the *cohort component approach*, where the population is aged from year to year and the components are organized according to age and sex cohorts. A cohort is a group of persons who experience a certain event in a specified period of time. For the calculation of age and sex estimates, birth cohorts (those persons born during the same year) by sex are used. Therefore the data required for the cohort component method include demographic events, such as deaths, immigration, emigration, that can be directly linked to persons belonging to the same birth cohorts by sex.

Chapter 9 describes the application of the cohort component approach in greater detail. The chapters on the separate components will detail the manner in which the components are organized by age and sex.

1.1.3 Levels of estimates

Updating population estimates between censuses entails the use of data from administrative files or surveys. The quality of population estimates therefore depends on the availability of a number of administrative data files that are provided to Statistics Canada by Canadian and foreign government departments. Since some components are not available until several months after the reference date, three kinds of postcensal estimates are produced: preliminary postcensal (PP), updated postcensal (PR) and final postcensal (PD). When all the components are preliminary, the estimate is described as preliminary postcensal. When they are all final, the estimate is referred to as final postcensal. Any other combination of levels is referred as updated postcensal estimates. The time lag between the reference date and the release date is three to four months for preliminary estimates and two to three years for final estimates.

1.2 Intercensal population estimates, Canada, provinces and territories

Intercensal estimates are estimates of population for reference dates between two censuses. They are produced following each census in order to reconcile previous postcensal estimates with the new census counts adjusted for net census undercoverage, thus assuring the internal consistency of the estimation system.

The production of intercensal estimates involves two basic steps:

- (1) the calculation of the error of closure; and
- (2) the linear distribution of the error of closure by intercensal year.

The error of closure is defined as the difference between the postcensal population estimates on Census Day and the population enumerated in that census (after adjustment for net census undercount). Assuming that the coverage studies that follow each enumeration are unbiased, the adjusted census figures are considered exact.

More specifically, the error of closure is calculated as:

Equation 1.2:
$$\varepsilon = P - \mathbf{P}$$

where

- ε = error of closure;
- P = postcensal population estimate; and
- \mathbf{P} = census population after adjustment for net census undercoverage (censal estimate).

The error of closure comes from two sources: differences in the amount of undercoverage or overcoverage in successive censuses and errors in the components of population growth over the intercensal period.

This can be calculated for any disaggregated group, or for any summation of such disaggregations up to and including the total population.

1.2.1 Provincial and territorial intercensal estimates of total population

For the production of intercensal estimates it is assumed that the error of closure is a linear function of the time elapsed since the previous census. The production of intercensal estimates of total population involves two steps: the calculation of the error of closure (ϵ) as in Equation 1.2, and the distribution of this error uniformly over the intercensal period by an arithmetic function.

Intercensal estimates of total population are obtained using the following formulae:

Equation 1.3:
$$IP_t = P_t - \epsilon$$

Equation 1.4:
$$IP_{t-1} = P_{t-1} - \frac{4}{5}\epsilon$$

Equation 1.5:
$$IP_{t-2} = P_{t-2} - \frac{3}{5}\epsilon$$

Equation 1.6:
$$IP_{t-3} = P_{t-3} - \frac{2}{5}\epsilon$$

Equation 1.7:
$$IP_{t-4} = P_{t-4} - \frac{1}{5}\epsilon$$

where

- t = the most recent census year;
- IP = intercensal population estimate;
- P = postcensal population estimate;
- ϵ = error of closure.

This same linear method applies also to the quarterly intercensal estimates of total population.

1.2.2 Provincial and territorial intercensal population estimates by age and sex

The error of closure for each sex and single year of age is the difference between the enumerated (after adjustment for net census undercount) and estimated populations, calculated using the same method as is applied to the total population. The production of the intercensal estimates by age and sex involves three steps:

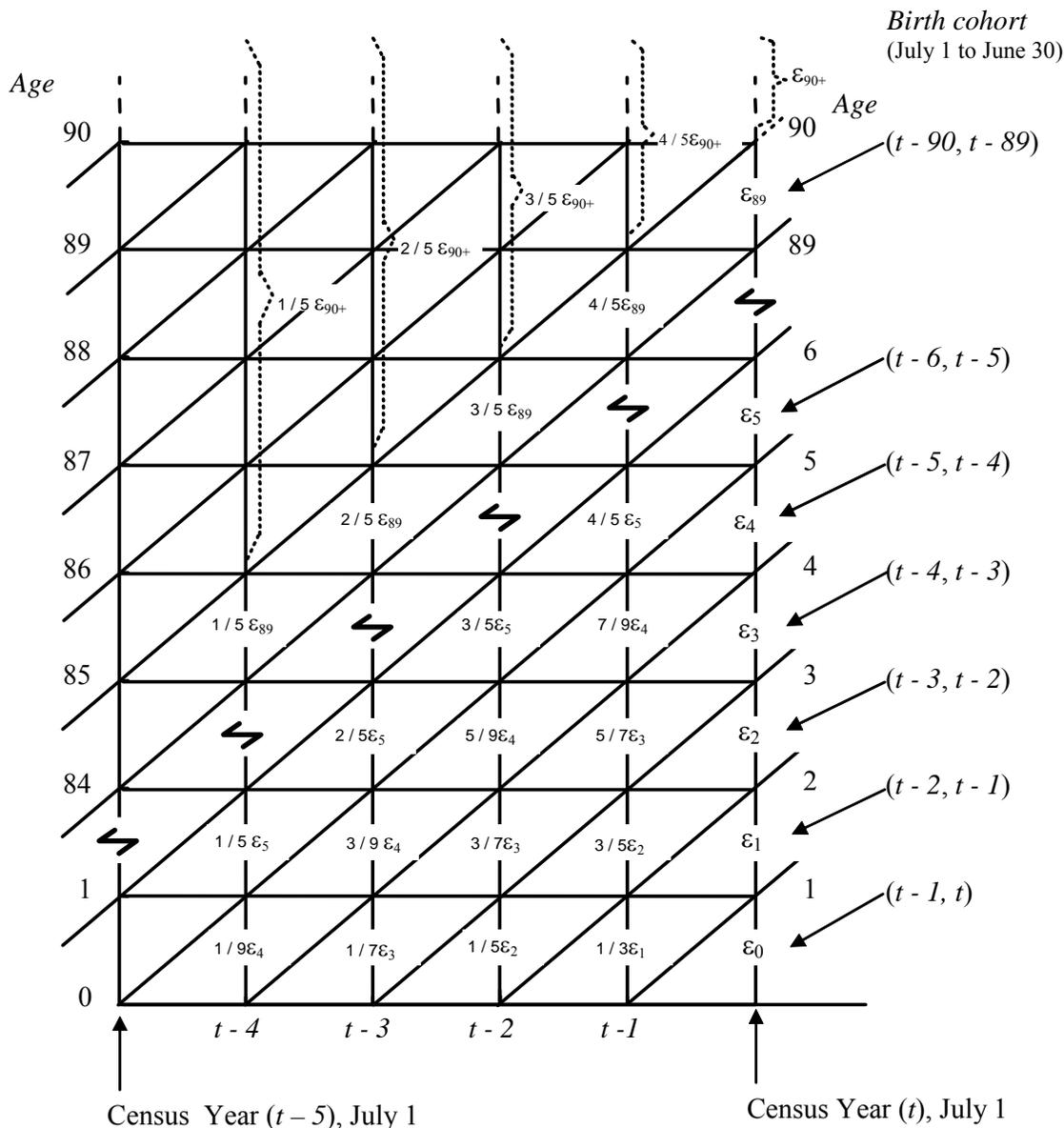
- (i) the calculation of the error of closure by age and sex;
- (ii) the distribution of this error; and
- (iii) a final adjustment to ensure consistency with total population figures estimated independently.

With the exception of ages between 0 and 4 years, and “90+”, the error of closure associated with each sex and single year of age is distributed linearly, as a function of the time elapsed since the previous census. Distributing the closure error between censuses following specific cohorts generates intercensal estimates. Figure 1.2 shows the method for distributing the error of closure.

For the 0-4 age group, the error of closure can only be distributed over the number of years elapsed since birth, but these children were not born at the time of the previous census, taken five years earlier. Accordingly, the error of closure for children aged 4 is distributed over nine half-year periods (4 ½ years), using 1/9, 3/9, 5/9, 7/9 and 9/9 for the first, second, third, fourth and fifth year from the year of the earlier census. For children aged 3, the error of closure is distributed using 1/7, 3/7, 5/7 and 7/7 over seven half-year periods, or 3 ½ years, and for those aged 2, using 1/5, 3/5 and 5/5. Finally for 1-year-old children, the error of closure is divided using the factors 1/3 and 3/3, covering a 1½-year period.

Figure 1.2
Method of distributing the error of closure by cohort

(ϵ_a = Closure error at age a)



Distribution of the error term for the 90+ age group must take the open-ended nature of the age interval into account. The error of closure for the 90+ age category at the reference date, t , relates to the cohort that was aged 85 at $(t - 5)$, 86 at $(t - 4)$, and so on. Error must be broken down for the period for the single years of age 86 through 89 and the 90+ age group. Error is therefore distributed in proportion to the number of years elapsed since the most recent census (i.e. 1/5, 2/5...) and by single years of age prorated over the population in the various age categories (86+ at $(t - 4)$, 87+ at $(t - 3)$, etc.) as reflected in the postcensal estimates.

For example, the error of closure (ε_{year}^{age}) for the year $(t - 3)$, at age 88, is calculated as:

Equation 1.8:

$$\varepsilon_{(t-3)}^{88} = \left(\frac{P_{(t-3)}^{88}}{P_{(t-3)}^{87+}} \right) \times \left(\frac{2}{5} \varepsilon^{90+} \right)$$

and for age 90+:

Equation 1.9:

$$\varepsilon_{(t-3)}^{90+} = \left(\frac{P_{(t-3)}^{90+}}{P_{(t-3)}^{87+}} \right) \times \left(\frac{2}{5} \varepsilon^{90+} \right)$$

Having determined the error of closure for each age and sex for the year $(t - 4)$ through to t , the intercensal estimates are obtained by subtracting this error from the postcensal estimate for the corresponding year. When summed, the series of intercensal estimates by age and sex, produced by taking the error of closure into account, differ slightly from the intercensal estimates of total population. These differences are then proportionally distributed among the age cohorts, accounting for the final adjustment made to the intercensal estimates at the provincial and territorial level.

Chapter 2

Base population (Censal estimates)

A *base population* is the population at the beginning of a period used as a reference or starting point for the estimation process. For postcensal estimates, the base population is the population enumerated in the most recent census, adjusted for net census undercoverage (also referred to as the “censal estimate”⁴). Postcensal estimates as of July 1 of a census year are obtained by the component method, using the most recent census of population adjusted to July 1 and for net census undercount. The base population for intercensal estimates is referred to as the “postcensal-estimated population”.

Since net census undercoverage (census undercoverage minus census overcoverage) is an important aspect of estimating population counts used in the population estimation program, this chapter focuses on the census population, adjusted for net undercoverage. It begins with a brief description of census collection procedures, followed by a discussion on studies used to provide estimates of census coverage error with a section on other errors associated with census counts, and concludes with procedures used for estimating net undercoverage.

2.1 Censal estimates as the base population

The census requires the participation of the entire population of Canada. Every five years, Statistics Canada conducts a census. The Census of Canada up to and including the 1966 Census have been conducted by interview. Starting from the 1971 Census, two collection methods have been used: self-enumeration and interview. The self-enumeration methodology is used in about 98% of households in the country except for some core areas of large urban areas, remote and northern areas, and most Indian reserves and settlements, where people are more difficult to enumerate. The census representative (CR) drops off a questionnaire, with instructions for an adult in the household to complete it for all persons living in the household on Census Day and to return it through the mail.

The remaining 2% of households are enumerated by personal interviews. With this method, the CR visits the household and completes the questionnaire during the interview.

Following the completion of collection by the CR, the work is checked by the CR’s supervisor, the Census Commissioner, and by a Quality Control Technician. Once the work is approved, the questionnaires and visitation records are forwarded to the data processing operations.

Details about census data collection and data processing procedures are described in the *Census Technical Reports*⁵ and in the *Census Handbook*⁶, published following each census.

4. The reference date for the censal estimate is the same as the corresponding census.

5. For the 2001 Census, refer to **Coverage, 2001 Census Technical Report (Reference Products 2001 Census)**, Catalogue No. 92-394-XIE, Statistics Canada.

6. For the 2001 Census, refer to **2001 Census Handbook (Reference)**, Catalogue No. 92-379-XIE, Statistics Canada.

The base populations in the population estimation program are derived from the quinquennial censuses between 1971 and 2001. The population universe of the 2001 Census includes the following groups:

- Canadian citizens (by birth or by naturalization) and landed immigrants living in Canada⁷;
- Canadian citizens (by birth or by naturalization) and landed immigrants who are abroad, either on a military base or attached to a diplomatic mission;
- Canadian citizens (by birth or by naturalization) and landed immigrants at sea or in port aboard merchant vessels under Canadian registry;
- persons with a usual place of residence in Canada who are claiming refugee status and members of their families living with them;
- persons with a usual place of residence in Canada who hold student authorizations (student visas or student permits) and members of their families living with them;
- persons with a usual place of residence in Canada who hold employment authorizations (or work permits) and members of their families living with them;
- persons with a usual place of residence in Canada who hold Minister's permits (including extensions) and members of their families living with them.

For census purposes, the last four groups in this list are referred to as “non-permanent residents” (NPRs). They are included in the population universe since the 1991 Census.

The 2001 Census population universe does not include foreign residents as is the case since the 1991 Census. Foreign residents are persons who belong to the following groups:

- government representatives of another country attached to the embassy, high commission, or other diplomatic body of that country in Canada, and members of their families living with them;
- members of the Armed Forces of another country who are stationed in Canada and members of their families living with them;
- residents of another country visiting Canada temporarily (for example, a foreign visitor on vacation or on business, with or without a visitor's permit).

The Canadian census uses the *de jure* method of enumeration, whereby persons are to be enumerated at their usual place of residence, even if they are temporarily away at the time of the census. Persons away from their usual place of residence and residing elsewhere in Canada are to be enumerated at their usual place of residence and are considered “temporary residents” at the other location (“temporary residents” should not be confused with “non-permanent residents”, which refers to the legal status of the person while in Canada). Persons without a usual place of residence are to be enumerated wherever they happen to be on Census Day.

Each base population for the population estimation program (P_t , where t = the census year) is adjusted as follows (unless otherwise noted, adjustments to the base population apply to both provincial, territorial and subprovincial levels):

7. The census enumerates the Canadian population at their usual place of residence but it also enumerates residents in collective dwellings and people not having a fixed address or living in unusual residential situations.

1. adjustment of the population for net undercoverage;
2. addition of independent estimates for incompletely enumerated Indian reserves in 1986, 1991, 1996 and 2001;
3. addition of estimates of non-permanent residents in 1971, 1976, 1981 and 1986. Since 1991, non-permanent residents have been included in the census universe; and,
4. estimation of the July 1 base by addition or subtraction of the components of growth between Census Day and June 30. At the subprovincial level, the estimate of the July 1 base population is obtained by applying the provincial age-sex distribution of the adjusted census to the subprovincial population estimate⁸.

2.2 Adjustment for net census undercoverage

Coverage errors are defined as errors caused by the miscounting of the population on Census Day. There are two types of coverage errors: undercoverage and overcoverage. Undercoverage occurs when a person who is a part of a census universe is missed by the census. Overcoverage, on the other hand, may occur in two ways. First, and most common, is the situation when a person who is part of a census universe is enumerated more than once. Second, a person outside the census universe, such as a foreign resident or a fictitious person may be erroneously enumerated.

A geographic error alone does not constitute a coverage error. That is, a person who is enumerated in the wrong geographic area does not constitute overcoverage for the area in which the person was enumerated and undercoverage for the area in which the person should have been.

Following each census, Statistics Canada undertakes coverage studies to measure coverage errors. Coverage studies provide undercoverage estimates for the 1991, 1996, and 2001 censuses at the provincial and territorial levels, and for the 1971, 1976, 1981 and 1986 censuses at the provincial level only. Estimates of overcoverage at the provincial and territorial levels are available only for the 1991, 1996 and 2001 censuses. Overcoverage for previous censuses was estimated by assuming that the overcoverage-to-undercoverage ratio for each census between 1971 and 1986 was the same as in 1991. The net census undercoverage for the Yukon Territory and the Northwest Territories prior to 1991 was estimated by assuming that the ratio between the net census undercoverage for each territory and the 10 provinces for each census between 1971 and 1986 was the same as in 1991.

For consistency, 1991 Census undercoverage and overcoverage were revised in 1998 to take into account the methodological improvements made in the 1996 Census coverage studies. This revision altered the net census undercoverage in all censuses between 1971 and 1986. Similarly, 1996 Census undercoverage and overcoverage were revised in 2003.

The following discussions on the procedures to estimate net census undercoverage are based on the 2001 Census coverage studies⁹.

-
8. The base population is also adjusted for early enumeration in 1991 and 1996 in parts of northern Quebec, Newfoundland and Labrador, the Yukon Territory, and the Northwest Territories. No adjustments for early enumeration were done in 2001.
 9. For details on coverage errors for the 2001 Census, refer to **Coverage, 2001 Census Technical Report (Reference Products 2001 Census)**, Catalogue No. 92-394-XIE, Statistics Canada.

2.2.1 Coverage Error Measurement Program

A number of changes have been made to the coverage studies for the 2001 Census as well as the methodology used for the calculation of net undercoverage by single year of age and sex. The 2001 Coverage Error Measurement Program focuses on the population and the dwelling universes. The following components of coverage error are measured:

- classification errors involving unoccupied private dwellings and census non-response dwellings;
- undercoverage and overcoverage of the population.

The 2001 Coverage Error Measurement Program consists of four studies:

- Dwelling Classification Study (DCS)
- Reverse Record Check (RRC)
- Automated Match Study (AMS)
- Collective Dwelling Study (CDS)

The DCS produces estimates of undercoverage arising from the incorrect classification of dwellings as unoccupied and from census non-response. The RRC measures undercoverage from all sources, including the undercoverage measured by the DCS. The RRC, AMS, and CDS all measure estimates of overcoverage. The AMS focuses on persons enumerated in more than one household within the same region (Atlantic, Quebec, Ontario, rest of Canada), while the CDS estimates overcoverage resulting from persons enumerated as usual residents in a collective dwelling who are also enumerated in a private dwelling. Although RRC is designed to measure overcoverage from all sources, overcoverage is removed from the RRC if it is detected in the AMS or the CDS.

The data resulting from these studies are used in the following ways:

- estimates from the DCS are included in the final census counts to account for this specific source of undercoverage;
- estimates from the RRC, the AMS, and the CDS are not included in the census counts. They are included in the base population for the population estimation program of Statistics Canada;
- information on the causes and characteristics of coverage errors is used in the planning of the next census to identify areas for sub-groups of the population where the level of coverage error is particularly high;
- supplementary information collected by the studies is used to evaluate the quality of selected census questions;
- the results serve to inform users about the nature and levels of coverage errors in the census so that they are better informed when drawing conclusions or making decisions based on census results.

Changes to the 1996 coverage studies are implemented to improve their quality and provide a more accurate measurement of coverage error in 2001. These improvements are described below.

- The Vacancy Check (VC) is replaced by the Dwelling Classification Study (DCS). This new study is used to correct census count before publication, for dwellings wrongly classified as

unoccupied and dwellings identified as occupied but the census is not able to obtain a contact, by collecting information on the characteristics of the occupants of those non-response dwellings. The DCS is more comprehensive and more precise than the old VC.

- Data for the RRC are collected by computer-assisted telephone interviewing (CATI). This collection method helps to improve data quality as it has the capability to do automated checks and to correct data instantly. It has also substantially reduced delays in submitting data to head office and in running additional checks on the data.
- The system that processes the addresses for the selected persons in the RRC is automated. At the same time, more sophisticated research tools, such as electronic telephone books and maps are used to improve the results. The automated system is more efficient as it is possible to process more addresses and significantly reduces the number still unresolved at the end of processing.
- The 2001 RRC produces estimates for all coverage errors, including those already corrected in the published census counts. Adjustments to the census counts based on the results of the DCS are removed, and the results of the RRC are used for adjusting undercoverage.
- The institutional component of the CDS (an overcoverage study) is dropped, and overcoverage estimates for this group are made as part of the RRC.

The methodology of each of the 2001 coverage error measurement studies are described below.

A. Dwelling Classification Study (DCS)

One of the potential sources of error in the census is the misclassification of dwellings. The misclassification of occupied dwellings as unoccupied results in the undercoverage of both persons and dwellings in the census counts, while the misclassification of unoccupied dwellings as occupied results in an overcount of persons and occupied dwellings. The latter misclassification of dwelling can occur when a dwelling was identified as occupied by a census representative (CR) but did not return a census questionnaire. This type of dwelling is referred to as a non-response dwelling. Moreover, the erroneous inclusion of marginal dwellings or dwellings under construction in the unoccupied dwellings classification results in the overcoverage of the housing stock.

The purpose of the Dwelling Classification Study (DCS) is to study these types of classification error. The DCS is an extension of the Vacancy Check which was used in previous censuses to re-examine dwellings which were classified as unoccupied by the census representatives. The study scope of DCS in 2001 is expanded to examine non-response dwellings as well as unoccupied dwellings.

The target population of the DCS is all non-response dwellings and all unoccupied dwellings on Census Day, excluding dwellings in collective enumeration areas (EAs), canvasser EAs, and Indian reserve EAs. These areas are excluded from the sampling frame because of cost and operational considerations. The sampling frame includes all self-enumeration EAs, which are then split into separate urban and rural frames for sample selection.

After all the data from the DCS have been captured and verified, the final step is the actual adjustment of the census databases. This is accomplished for the unoccupied dwellings by first producing a profile of estimated number of misclassified dwellings with the household size for both urban and rural areas, using the type of dwelling and the number of persons missed because of the

misclassification. These national profiles are then used to create estimates of the number of misclassified dwellings by number of persons in the household, type of dwelling, and rural/urban parts at the province and territory level. On the basis of these estimates, enumerated households with the same characteristics (number of persons, type of private dwelling) are selected at random, and their weights in the census are increased by one unit. For each household selected, the weight of one unoccupied dwelling from the same EA is set to zero so that the total number of dwellings will remain the same.

The number of persons living in non-response dwellings is adjusted in the census databases in the following way. A household size entered on the questionnaire by the CR in the field for a non-response dwelling is accepted as correct. A profile of non-response dwellings that gives the estimated number of dwellings for each household size (from zero to six persons) for each geographic area is generated by the DCS. A second distribution is created from the observed census data, and a misclassification vector which is the difference between these two distributions is also created. Non-response households are then randomly selected from those available in each geographic area, and new household sizes are imputed for them until the misclassification vector is equal to zero, or until the distribution of household sizes for non-response households in the census equals the distribution of non-response household sizes from the DCS.

B. Automated Match Study (AMS)

Overcoverage in a census occurs when a unit is enumerated more than once or when a unit is enumerated when it should not be. Persons who are included more than once cause most of the overcoverage. The level of overcoverage in the 2001 Census is measured by three studies: the Automated Match Study (AMS), the Collective Dwelling Study (CDS), and the Reverse Record Check (RRC).

The AMS is designed to detect and estimate overcoverage due to duplication between different private households by searching the census database for pairs of households containing persons that have the same sex and full date of birth characteristics and are located in the same geographic region (Atlantic, Quebec, Ontario, Western Canada, which includes British Columbia and the territories). In 2001, a new operation is added to find pairs of households in the Ottawa-Hull Census Metropolitan Area, with one household in Quebec and the other in Ontario. This covers a large urban area that crosses two different geographic regions where a lot of mobility occurs.

Two sets of programs are in place in the AMS. The first set identifies pairs of census households in the same region with at least two exact person matches, whereas the second set identifies pairs in the same Federal Electoral District (FED) with only one exact person match. Matching weights are assigned to the pairs of households which are then stratified. The pairs of households within the strata with the highest similarity weights are automatically classified as overcoverage. Within the other ones, a sample of pairs of households is selected and their census documents were reviewed. If the same person is on both questionnaires, overcoverage has occurred. Since the Census Day address is unknown, the weight of the sample unit is split between the households where overcoverage occurred.

C. Collective Dwelling Study (CDS)

This study covers all persons who are enumerated as usual residents in a non-institutional collective dwelling and measures overcoverage which occurs when a respondent completes two census questionnaires at different dwellings. A collective dwelling is a dwelling of commercial, institutional, or communal nature. Collective dwellings are of two types: institutional dwellings such as hospitals and jails, and non-institutional dwellings such as hotels, YMCAs, and school residences (Statistics Canada, 1999). Contrary to 1996, the 2001 Collective Dwelling Study (CDS) does not produce estimates for institutional collective dwellings. The source of overcoverage for these is now covered by the Reverse Record Check.

During the enumeration of collective dwellings on Census Day, usual residents of non-institutional collective dwellings completed an individual census form on which they were asked to report an alternate address where they could have been enumerated. These addresses provide a sampling frame for the Collective Dwelling Study (CDS).

The sample is allocated to each province and territory proportionally to the square root of its non-institutional collective usual residents with a minimum of 75 persons per province and territory. Then a sample is allocated by dwelling type within each province and territory in proportion to its population of usual residents.

For each individual in the sample, the questionnaire completed at the alternate address is verified to determine if the individual is enumerated at a private dwelling a second time. If the individual is found on the questionnaire, overcoverage has occurred.

D. Reverse Record Check (RRC)

Population and household undercoverage is generally considered as one of the most important sources of error affecting census data. Undercoverage causes the census figures to underestimate the true population and household counts, while overcoverage results in the opposite or overestimation. These two sources of error can also distort the distribution of the population and household characteristics estimated from the census data if the missed and overcounted persons do not have the same characteristics as the estimated individuals.

The RRC measures both gross undercoverage and gross overcoverage. Its results are combined with the findings of the other coverage studies to calculate net undercoverage. In 1996, for the first time, the Private Dwelling Study carried out for 1991 to measure overcoverage, is incorporated into the RRC.

The RRC sample consists of persons who should have been enumerated in the census and was selected from sources independent of the census. Shortly after the census field operations are finished, tracing operations are undertaken to contact and interview the persons in the sample and find out if they were part of the census target population, where they were living on Census Day and obtain many other addresses where they could have lived and been enumerated. Subsequently, census documents are searched to determine whether the Selected Person (SP) had been enumerated and, if so, whether they had been enumerated more than once or erroneously. Deceased persons, for example, who were enumerated, are considered to have been enumerated erroneously.

The RRC sample is selected from six sources or sampling frames. The first five sampling frames are used to estimate undercoverage and overcoverage in the ten provinces, whereas estimates for the three territories are calculated on the basis of samples from the sixth frame. The six sampling frames of the 2001 RRC are:

1. 1996 Census: all persons enumerated in the 1996 Census;
2. 1996 Missed: all persons from the 1996 RRC sample who were classified as “not enumerated”;
3. Births: all children born between May 14, 1996 and May 14, 2001;
4. Immigrants: all landed immigrants who arrived in Canada between May 14, 1996 and May 14, 2001;
5. Non-permanent residents: all persons holding employment or student authorizations or Minister’s permits, including extensions, and persons claiming refugee status who were in Canada on Census Day (May 15, 2001); and
6. Health care files: all persons listed in the health insurance files of the Yukon Territory, Northwest Territories, and Nunavut. These files are updated by the territories to reflect the population around Census Day.

Sample design and selection are carried out independently within each frame. Sampling rates from each of these frames depends on the size of the frame, the expected rate of tracing and interviewing, and the expected gross undercoverage and gross overcoverage.

A problem that exists with using multiple frames is the possibility that persons may be listed on more than one frame. For example, a person in the Immigrants frame may have been in Canada on a work permit at the time of the second-most recent census and thus would have been enumerable for that census. Consequently, this person would then be in the Immigrants frame and in the Census frame if he or she was enumerated, or represented in the Missed frame if he or she was not enumerated. All potential cases of frame overlap must be identified to avoid double-counting.

The purpose of the various RRC operations is to classify each SP as one of the following:

1. enumerated once in the census (excluding persons added by the DCS);
2. enumerated more than once in the census;
3. missed in the census;
4. died before the census;
5. emigrated before the census;
6. temporarily abroad at the time of the census;
7. out of scope, that is, the SP should not be included in the census (for example, babies born after census day, permit holders and refugee status claimants who were no longer in Canada on Census Day) or covered by more than one frame;
8. not identified: an identifier such as name, date of birth or sex is missing, resulting in insufficient information to initiate or validate the tracing process;
9. not traced: the SP was not contacted and interviewed to determine whether he or she belonged to the census target population and, if so, what his or her usual residence was on Census Day;
or
10. not classified: it is determined through contact that the SP was part of the target population but there was insufficient information to assign a final classification because the addresses were too vague.

In 2001, the RRC classifies persons added by the DCS as missed and provides an estimate for them because the RRC offer more accurate estimates than the DCS. Thus, the persons added to the census by the DCS are netted out for estimation purposes. In 1996, the RRC did not estimate the persons added by the Vacancy Check (that was replaced by the DCS in 2001).

The classification of each SP is achieved by two operations: collection and address searching. The purpose of collection is to establish the SP's address and status on Census Day as enumerable or not enumerable. This is accomplished through conducting telephone interviews with each SP. All addresses where the SP could have been enumerated on Census Day are also collected. Collection begins with sending sampling frame data and other data such as previous phone numbers and addresses to Statistics Canada's Regional Offices. Interviewers then make every effort to trace the SP and conduct a telephone interview. In 2001, for the first time, the interviewers use Computer Assisted Telephone Interviews system. This allows the daily transfer of information from the regional offices to head office with no further data capture requirements.

Address searching consists of checking the census database and images of the census documents, such as visitation records and questionnaires from the most recent census for each address provided by the SP to determine whether the SP had been enumerated there. In 2001, the images of the census questionnaires are consulted through an imagery system replacing the need to consult the actual questionnaires. This is much more efficient and reduces the management of documents substantially.

An operation introduced in 1996 and known as the "monster-match" was used in the 2001 Census. This operation helps to classify SPs as enumerated at addresses not identified during tracing or at unsearchable addresses. It consists of matching the SP's household against the 2001 Census database, using the information on date of birth and sex, and identifying all households in the database with two or more matching members. The images of these household questionnaires are subsequently consulted to verify if the SPs were enumerated.

Prior to final classification, additional searching is done for all SPs that have not yet been found enumerated. Efforts are made to identify addresses using electronic telephone books or tax data.

Finally, all SPs are classified into one of the categories mentioned above, and the creation of a final database can begin. Data from the interviews provide the basis for classifying SPs as contacted, emigrated, temporarily abroad, or not traced. Two sources of data are required to classify an SP as deceased. On the basis of the address search results, enumerable SPs are classified as enumerated once, enumerated more than once, or missed.

The process of creating the final database consists of five main steps:

1. extraction and edit of selected data from the RRC questionnaire and the sample control file;
2. imputation of item non-response for the RRC questionnaire;
3. incorporation of address search and classification results;
4. processing for total non-response and other weight adjustments; and
5. calculation of final estimates of undercoverage, overcoverage and their standard errors.

All data in the RRC questionnaires are captured by the interviewers. In the cases where data for one or more questions are missing or inaccurate, an attempt is made to obtain the required information from other sources such as sample frames and administrative records. The results of address processing and SP classification are then added to the final database.

The first step in producing estimates from the final classification of the RRC sample is to adjust the sampling weights of respondents to account for all types of non-respondents (“not identified”, “not traced” and “not classified”). The adjustment consists of redistributing the weight of groups of non-respondents among groups of respondents assumed to have a similar probability of being missed in the census. A major criterion in forming the groups is the SP’s potential mobility, since it is reasonable to assume that a person changing residence is more likely to be missed than a person who is not mobile. A person that requires the collected information to be classified is defined as mobile, thus non mobile persons are classified with the information available before collection.

The final step is a post-stratification, or “weight adjustment,” which ensures compatibility with known totals in the sample frame populations. First, the slight incomplete coverage of the Health Care Files of the territories is taken into account. Second, an adjustment is made for non-permanent residents using more precise figures. Finally, age and sex differences between the census database and the RRC database are addressed.

An extra weight adjustment is carried out for those SPs classified as enumerated more than once. In some cases, the addresses provided by the SP are too vague to identify a particular address. The weight adjustment accounts for this type of non-response.

Although the RRC measures the same types of overcoverage as the Automated Match Study (AMS) and the Collective Dwelling Study (CDS), it does so with much less accuracy. The weights of the RRC overcoverage cases are higher than the weights for the AMS or CDS overcoverage cases. Consequently, only cases of overcoverage that cannot be identified by the two other studies are considered by the RRC.

2.3 Calculating net undercoverage

The net coverage error is defined as:

Equation 2.1:
$$NU = \mathbf{P} - \mathbf{C}$$

where

- NU = net coverage error;
- P** = censal estimate (total number of persons who should have been enumerated in the census);
- C** = census count.

The censal estimate **P** is defined as:

Equation 2.2:
$$\mathbf{P} = TU + EN$$

where

- TU = total undercoverage (number of persons missed in the census whether included or not by the DCS);
- EN = number of persons who were enumerated at least once (excluding persons added by the DCS).

The undercoverage excluding persons added by the DCS is defined as:

$$\text{Equation 2.3:} \quad U = TU - \text{DCS}$$

where

U = undercoverage (number of persons missed in the census and not included by the DCS);
 DCS = number of persons included by the DCS.

Therefore:

$$\begin{aligned} \text{Equation 2.4:} \quad \text{NU} &= (TU + \text{EN}) - \mathbf{C} \\ &= TU - (\mathbf{C} - \text{EN}) \\ &= TU - \text{DCS} - (\mathbf{C} - \text{EN} - \text{DCS}) \\ &= TU - \text{DCS} - \mathbf{O} \\ &= U - \mathbf{O} \end{aligned}$$

where the term ($\mathbf{O} = \mathbf{C} - \text{EN} - \text{DCS}$) is defined as overcoverage. This error results not only from counting units more than once, but also from counting units that are outside the universe in question.

Undercoverage, overcoverage, and the net of the two, net undercoverage, are most usefully expressed as rates of the population that should have been enumerated in the census. The rates of coverage error are:

$$\text{Equation 2.5:} \quad \text{Undercoverage rate:} \quad R_U = \frac{U}{\mathbf{P}}$$

$$\text{Equation 2.6:} \quad \text{Overcoverage rate:} \quad R_O = \frac{\mathbf{O}}{\mathbf{P}}$$

$$\text{Equation 2.7:} \quad \text{Net undercoverage rate:} \quad R_{NU} = \frac{\text{NU}}{\mathbf{P}}$$

or:

$$\begin{aligned} \text{Equation 2.8:} \quad \text{Net undercoverage rate:} \quad R_{NU} &= \frac{(U - \mathbf{O})}{\mathbf{P}} \\ &= R_U - R_O \end{aligned}$$

The population undercoverage rate \hat{R}_U , indicating the proportion of persons missed by the census as a proportion of the total number of persons who should have been enumerated, is calculated as follows:

$$\text{Equation 2.9:} \quad \hat{R}_U = \frac{\hat{M} - \hat{D}}{\mathbf{C} + (\hat{M} - \hat{D}) - \hat{O}}$$

where

- \hat{M} = estimate of the number of persons not enumerated at their usual place of residence, as obtained from the RRC;
- \mathbf{C} = census count;
- \hat{D} = estimate of the number of persons not enumerated because they occupied dwellings classified by the census representative as unoccupied, as obtained from the DCS;
- \hat{O} = estimate of the number of persons enumerated more than once or in error

where

$$\hat{O} = \hat{O}_{AMS} + \hat{O}_{CDS} + \hat{O}_{RRC}$$

where

- \hat{O}_{AMS} = estimate of overcoverage measured by the Automated Match Study;
- \hat{O}_{CDS} = estimate of overcoverage measured by the Collective Dwelling Study;
- \hat{O}_{RRC} = estimate of overcoverage measured by the Reverse Record Check, net of the overcoverage included in the Automated Match Study and the Collective Dwelling Study.

The denominator of \hat{R}_U estimates the number of persons who should have been enumerated as:

- the total number of persons actually enumerated ($\mathbf{C} - \hat{D}$); plus
- the number of persons who should have been enumerated but were not (\hat{M}); minus
- the number of persons enumerated who should not have been enumerated (\hat{O}).

\hat{D} is subtracted from \mathbf{C} because the published census count includes an estimate of the number of persons who were erroneously excluded because they occupied dwellings misclassified as unoccupied.

The population overcoverage rate \hat{R}_O is calculated as follows:

Equation 2.10:

$$\hat{R}_O = \frac{\hat{O}}{\mathbf{C} + (\hat{M} - \hat{D}) - \hat{O}}$$

The population net undercoverage rate \hat{R}_{NU} is calculated as follows:

Equation 2.11:

$$\begin{aligned} \hat{R}_{NU} &= \hat{R}_U - \hat{R}_O \\ &= \frac{(\hat{M} - \hat{D}) - \hat{O}}{\mathbf{C} + (\hat{M} - \hat{D}) - \hat{O}} \end{aligned}$$

A positive net undercoverage rate indicates that undercoverage is larger than overcoverage, whereas a negative net undercoverage rate indicates that undercoverage is smaller than overcoverage. In most cases, undercoverage is larger than overcoverage. Thus, most net undercoverage rates are positive.

2.4 Adjustments for non-enumerated Indian reserves and settlements

Enumeration is sometimes not permitted on some Indian reserves and settlements or it is interrupted before it can be completed. These areas, a total of 30 in the 2001 Census, are called “incompletely enumerated Indian reserves and Indian settlements”. Census data for these areas are not available and therefore have not been included in any census tabulations.

These areas present unique problems for the coverage studies and for the population estimation program. The RRC target population is the same as the census; however the population that is actually covered by the RRC does not include those residents for which the 2001 Census did not attempt to collect any data. For the purposes of the RRC, these residents would be considered to be out-of-scope. Yet, the population estimation program requires an estimate of the population living in these areas. Since neither the census nor the RRC is in a position to produce such an estimate, another approach is used to estimate this population independently. These estimates are subsequently added to the adjusted census figures for the base population.

A two-step estimation model is developed to estimate the 2001 incompletely enumerated Indian reserve and Indian settlement population. First, a simple linear regression is built to predict the census count in 2001. Then, in order to be consistent with the RRC results, an adjustment is made to the estimated “census” count to account for net undercoverage that all census counts are subject to.

The linear regression is constructed using all Indian reserves and settlements that were completely enumerated in both the 1996 and 2001 censuses. The basic model assumes a linear growth from 1996 to 2001 for all provinces. However, for the intercept and the regression parameters, separate estimates are produced for each province for this model. The output of the model gives the estimated size of non-enumerated Indian reserves and settlements in the 2001 Census. These census counts, in order to be consistent with all population estimates from the RRC, have to be adjusted for net undercoverage. The net missed persons are estimated by calculating the net undercoverage rate for all completely enumerated Indian reserves and settlements in each province. This provincial rate is then applied to the estimated “census” count of all the incompletely enumerated Indian reserves and settlements in the province.

The estimated “census” count and the “estimated net missed persons” count in each reserve or settlement are then summed to create an “estimated” population for the incompletely enumerated Indian reserve or settlement.

2.5 Estimates of net undercoverage by single year of age and sex

2.5.1 Provinces and territories

The RRC is designed to produce reliable direct estimates for large areas, such as provinces and for large domain, such as age-sex combinations at the national level. However, the population estimation program requires estimates of net undercoverage by single year of age and sex for each province and territory. These are derived using various procedures which include:

1. direct survey to produce estimates of total net undercount by province and territory (provincial and territorial marginal totals);
2. a calibrated curve smoothing method to produce the national estimates of net undercount by age and sex (national age-sex marginal total);
3. Empirical Bayes regression model to create estimates of net undercount at the provincial and territorial level, for broad age and sex categories;
4. synthetic model to generate estimates of net undercount by single year of age;
5. raking ratio procedure to ensure the consistency of estimates of net undercount by single year of age and sex with the provincial/territorial totals and national age-sex totals.

Marginal totals – Provincial and territorial margin

The estimation procedures require two marginal totals: the provincial/territorial estimate of net undercount and the national total of net undercount by single year of age for each sex. The provincial and territorial estimates of net undercount can be found in the official population release on census coverage studies. These are assumed to be correct and all domain estimates will be constrained to these totals.

Marginal total – National age-sex margin

The direct survey estimates of national net undercount by age and sex require some smoothing. The direct estimates of net undercoverage rates by single year of age (for each sex) from the coverage studies are used as the input data. However, unless the net undercoverage rates between consecutive ages are smoothed, these estimates cannot be used as marginal totals. There are a variety of methods that can be used to smooth the direct estimates; in 2001 a cubic smoothing spline was introduced. This approach implicitly assumes that consecutive age groups have relatively similar undercoverage rates and change between ages follows a smooth function.

The model assumes that the true undercoverage rates - defined as the ratio of net missed persons over the total population - are described by a smooth function of age. This approach is the continuous version of the discrete Whittaker-Henderson Graduation Method used in the 1996 population estimates (Gambino and Dick, 2000).

The RRC publishes the undercoverage rates by selected broad age groups: 0 to 19, 20 to 29, 30 to 44, 45 to 69 and 70 and over. The smoothing spline will produce different estimates for these groups. A common procedure to ensure that small domain estimates are in agreement with higher level aggregated data is to calibrate the estimates to the totals for the broader age groups (Ramsay, 2000).

Small domain estimation – Broad age groups

Detailed estimates of net undercount by single year of age and sex within each province/territory are handled in a two-step procedure. First, an Empirical Bayes regression model is used to derive broad age group and sex estimates of undercount at the provincial and territorial level, and then a synthetic estimate is created for the single years within the broad age groups.

The objective of modelling the small domain estimates of undercount is to produce a series of estimates with a smaller mean square error than the direct estimates. However, as opposed to the direct survey estimate which is design unbiased, the modelling approach will introduce a bias for each estimate. Thus modelling the small domain estimates of undercount involves a trade-off between reducing the variance of each estimate and the bias introduced through the modelling process. One approach to ensuring that the more reliable direct survey estimates are used is to introduce an Empirical Bayes model similar to Fay and Herriot (1979). This procedure creates an estimate of undercount that is a combination of a model estimate and the direct survey estimate weighted by their respective variances. Note that since the individual sampling variances are used in the estimation, a more precise direct estimate would contribute much more to the final Empirical Bayes estimate than a direct estimate with low precision. This ensures that the model does not dominate estimates that are already considered reliable.

The Empirical Bayes regression model starts with the direct survey estimates of adjustment factors (ratio of true population to census population) with variances for four broad age groups (0 to 19, 20 to 29, 30 to 44 and 45 and over) and sex for each province and territory. This approach is best described as a two-stage regression model: a sample model that describes the basic relationship between the estimated adjustment factors and the true adjustment factors, and a regression model that describes how the true adjustment factors vary with a set of underlying variables.

The regression coefficients are estimated using a weighted least squares method which ensures that direct estimates from the large provinces with small standard errors are respected more than direct estimates from the small provinces with large standard errors. Since the sample variance estimates are highly unstable, they are first stabilized using a generalized variance function. The estimated variances are then used and assumed to be without error in the subsequent step (for details, refer to Dick, 1995).

Small domain estimation – Synthetic expansion

The population estimation program requires estimates of net undercount by single year of age and sex for each province/territory. The Empirical Bayes estimates, discussed above, can produce estimates for broad age groups but it cannot work effectively with more detailed age categories due to a lack of sample. Too many domains would produce an estimate of zero since no sample was in the domain. To meet the requirements of the population estimation program, a synthetic estimate must be introduced.

Estimated adjustment factors for four broad age groups and sex in each province/territory are produced as a result of the Empirical Bayes model. A synthetic model is then used to expand the estimates from the broad age groups to single ages. This model essentially assumes that a constant net undercoverage rate within the four broad age groups is to be maintained. The adjustment factors (\hat{F}_{jk})

from the Empirical Bayes model allow for an easy synthetic estimate of missed persons for any single age a , within the k age group for area j using the following equation:

$$\text{Equation 2.12:} \quad \hat{M}_{jk_a} = C_{jk_a} \times (\hat{F}_{jk} - 1)$$

where

\hat{M}_{jk_a} = net number of persons missed in single year of age a in province/territory j and broad age group k ;

C_{jk_a} = number of persons counted in the census for single year of age a in province/territory j and broad age group k ;

\hat{F}_{jk} = adjustment factor produced by Empirical Bayes model by broad age group k .

Consistency adjustment

The small domain estimation will not be consistent with the marginal totals discussed above. Hence a raking ratio adjustment is used on the small domain estimates to ensure consistency with both the provincial totals and the national age-sex totals.

This procedure organizes the estimates of missed persons into a matrix with the single year of age estimates as the row and the province/territory estimates as the columns. The fixed marginal totals are then used for the single ages at the national levels and for the provincial totals. The Empirical Bayes estimates of the adjustment factors are then used to generate the synthetic estimates of missed persons for each province/territory. These estimates are then alternatively adjusted so that they sum to the row and column totals. Convergence is usually reached in about three iterations (for details, refer to Dick, 1995).

These estimates are then used as the small domain estimate of missed persons. By adding them to the census counts, the population estimation program can create a base population for generating the population estimates.

2.5.2 Subprovincial areas

Base populations for census metropolitan areas and census divisions are obtained by applying the corresponding provincial/territorial net undercount rates, available by age and sex. This synthetic estimate assumes that within a province/territory and for a single year of age there is a constant net undercoverage rate. For example, in British Columbia a 20 year old male would be assumed to be missed at the same rate across the entire province. Late enumeration and non-enumerated Indian reserves and settlements were adjusted by adding the provincial/territorial estimates to the appropriate geographic regions. All figures sum to provincial/territorial and national totals.

2.6 Estimates of net undercoverage by marital status (and age and sex)

2.6.1 Provinces and territories

At this point, estimates of net undercoverage are available by single year of age and sex for each province and territory. However, the population estimation program requires estimates of provincial/territorial net undercoverage by age, sex and marital status.

The basic requirement of the estimation procedure is for all the estimates produced by age, sex and legal marital status to respect the higher level estimates. Two marginal totals are used to generate the detailed estimates: age totals in each province/territory for each sex, and the RRC estimates of the net missed persons in each province/territory by marital status (single, married (or separated), divorced and widowed). The estimates of gross undercoverage are raked to ensure that all estimates can be aggregated to the two marginal totals.

In order to avoid extreme estimates of net undercoverage rates, two assumptions are established. First, it is assumed that all persons under the age of 25 that were missed in the census are single. Second, it is assumed that persons who are widowed could not be missed by the census unless they are at least 50 years of age.

Chapter 3

Births and deaths

Births and deaths have been recorded on a regular basis since 1921. Since the registration of births and deaths is a legal requirement in Canada, these data are readily available and are generally of very high quality¹⁰.

Live births are added to and deaths, excluding stillborns, are subtracted from the base population (along with the migration components) to derive the postcensal population estimates. Births are commonly used to mean “live births” and deaths are also used to mean “deaths excluding stillborns”. This chapter presents information related to the data sources for births and deaths, as well as the methods of estimation for preliminary estimates.

3.1 Data sources and relevant concepts

Information on births and deaths is obtained from the Vital Statistics databases maintained by Health Statistics Division of Statistics Canada. The databases are created in collaboration with all provinces and territories as they are responsible for the registration of vital events within their respective jurisdictions. Provincial and territorial Vital Statistics Acts (or equivalent legislation) render compulsory the registration of all live births and deaths within their jurisdictions. The central Vital Statistics Registry in each province and territory provides data from birth registrations and death registrations to Statistics Canada. The data are available at national, provincial and territorial, census division (CD), and census subdivision (CSD) levels¹¹.

The actual population of the birth database includes births to Canadian women, births to landed immigrant women and births to non-permanent resident women in Canada. It also includes births to Canadian women in some American states. The actual population of the death database is deaths of Canadian residents, deaths of landed immigrants and deaths of non-permanent residents in Canada. Deaths of Canadian residents in some American states are also included. Births and deaths of foreign residents are excluded.

The excess number of births over deaths during a given period is known as the “natural increase” of a population. If the number of births during a particular reference period is less than the number of

10. Undercoverage of births is thought to be minimal, but it can occur because of late registration; some births to Canadian resident women occurring outside Canada; non-registration of births (e.g., such as in Ontario), and for infants who die within days of the birth. Undercoverage of deaths is also thought to be minimal, but it can occur because of late registration; delayed or missing registrations with unidentified bodies, or for some Canadians who die outside of Canada; and deaths of Canadians who died overseas while serving in the Armed Forces are not included in the Statistics Canada databases because they are not registered by the provinces and territories. Overcoverage of births and deaths is not measured and deemed insignificant.

11. Data for births and deaths from the Vital Statistics databases are not readily available for Census Metropolitan Areas (CMAs), Demography Division derives these counts using data at lower levels of geography.

deaths during the same period, the natural increase will be negative. The rate of natural increase has been decreasing steadily since the 1990s, but it seems to have stabilized since 2000 (see Table 3.1).

Table 3.1
Births, deaths, natural increase, Canada, 1971/1972 to 2005/2006

Period (July 1 to June 30)	Births	Deaths	Natural increase	Year to year change	Natural increase/ Population (mid period) x 100
1971/1972	351,256	159,533	191,723	...	0.87
1972/1973	345,815	162,618	183,197	-8,526	0.82
1973/1974	342,446	166,284	176,162	-7,035	0.78
1974/1975	355,960	168,751	187,209	11,047	0.81
1975/1976	364,278	166,428	197,850	10,641	0.85
1976/1977	357,850	165,747	192,103	-5,747	0.81
1977/1978	359,793	169,030	190,763	-1,340	0.80
1978/1979	362,432	165,805	196,627	5,864	0.82
1979/1980	367,286	171,460	195,826	-801	0.80
1980/1981	372,139	170,535	201,604	5,778	0.82
1981/1982	372,472	172,352	200,120	-1,484	0.80
1982/1983	373,594	176,522	197,072	-3,048	0.78
1983/1984	374,533	174,159	200,374	3,302	0.79
1984/1985	376,265	179,085	197,180	-3,194	0.77
1985/1986	375,381	183,353	192,028	-5,152	0.74
1986/1987	373,022	182,599	190,423	-1,605	0.72
1987/1988	370,033	189,917	180,116	-10,307	0.68
1988/1989	384,035	188,408	195,627	15,511	0.72
1989/1990	403,280	192,608	210,672	15,045	0.77
1990/1991	402,929	192,439	210,490	-182	0.76
1991/1992	403,107	196,967	206,140	-4,350	0.73
1992/1993	392,181	201,808	190,373	-15,767	0.67
1993/1994	386,159	206,464	179,695	-10,678	0.62
1994/1995	381,998	209,389	172,609	-7,086	0.59
1995/1996	372,453	209,766	162,687	-9,922	0.55
1996/1997	357,313	217,221	140,092	-22,595	0.47
1997/1998	345,123	217,688	127,435	-12,657	0.42
1998/1999	338,295	217,632	120,663	-6,772	0.40
1999/2000	336,912	217,229	119,683	-980	0.39
2000/2001	327,107	219,114	107,993	-11,690	0.35
2001/2002	328,155	220,494	107,661	-332	0.35
2002/2003	330,523	223,905	106,618	-1,043	0.34
2003/2004	337,762	230,092	107,670	1,052	0.34
2004/2005	338,894	233,749	105,145	-2,525	0.33
2005/2006	343,517	234,914	108,603	3,458	0.33

... not applicable

Note: The numbers of births are updated for 2004/2005. The numbers of deaths are updated for 2003/2004 and 2004/2005. The numbers of births and deaths are preliminary for 2005/2006.

Source: Statistics Canada, Demography Division.

3.2 Birth and death estimates, Canada, provinces and territories

3.2.1 Levels of estimates

Vital statistics are not available for the current population estimates; births and deaths are estimated using rates. The number of births is estimated using fertility rates by mother's age. The number of deaths is estimated using mortality rates by age and sex. These methods are used to calculate preliminary quarterly estimates. Updated fertility and mortality rates are used to revise the preliminary estimates as it will take two years to get final annual vital estimates of births and deaths. Annually, as information from vital statistics is updated to finalize past estimates, new fertility and death rates are calculated. These new rates are used to revise all non-final estimates. Thus, the quarterly preliminary estimates may be estimated from rates that are up to three years old while annual preliminary estimates are from rates that are two years old. The fertility and death rates are revised every year.

3.2.2 Final estimates

Because of the completeness of the Vital Statistics databases, very few adjustments are required to the data received from Health Statistics Division (HSD) for demographic counts. When both place of occurrence and place of residence are available on the data files, the place of residence is the geography of choice for reporting the event for population estimates, as with vital statistics health indicators. When a mother has given birth out-of-province (i.e., a province other than her place of residence), the birth will be counted in the mother's province of residence. Similarly, in the event of the death of a person outside of the deceased province of residence, the death is counted in the province of residence. This is because the birth or death affects the population of residence rather than that of where the event occurred.

Using the component approach, between the date of the base population and the reference date of the estimate, the number of births is simply added and the number of deaths subtracted from the base population.

3.2.3 Preliminary estimates

Complete data on births and deaths are not available at the time the preliminary estimates are prepared and, therefore they have to be estimated. With the exception of Quebec and British Columbia, where preliminary birth and death estimates are provided by provincial statistical agencies, estimates of births and deaths are derived from the fertility and death rates from the most recent reference period for which final estimates are available.

The following formulae are used to obtain preliminary estimates of births and deaths at the national and provincial/territorial levels on a quarterly basis:

For births:

By quarter for each province and territory, by age group of mother at the beginning of the quarter:

Equation 3.1:

$$B_{(t,t+1)}^{a_M} = \left(\frac{f B^{a_M}}{f P^{a_F}} \right) \times P_t^{a_F}$$

where

- $(t, t + 1)$ = interval between the date of the previous reference period, time t , and the reference date of the estimate, time $t + I$;
- $B_{(t,t+1)}^{aM}$ = estimate of number of births from mother M age a between time t and $t + I$;
- ${}^f B^{aM}$ = number of births from mother M age a during the corresponding quarter for which final estimates f are available;
- ${}^f P^{aF}$ = population estimate of females F age a at the beginning of the corresponding quarter for which final birth estimates f are available;
- P_t^{aF} = population estimate of females F age a at the beginning of the quarter of estimation (time t).

And then, the estimate of total number of births:

Equation 3.2:
$$B_{(t,t+1)} = \sum_a B_{(t,t+1)}^{aM}$$

For deaths:

By quarter for each province and territory, by age at the beginning of the quarter, for each sex:

Equation 3.3:
$$D_{(t,t+1)}^{a,s} = \left(\frac{{}^f D^{a,s}}{{}^f P^{a,s}} \right) \times P_t^{a,s}$$

where

- $(t, t + 1)$ = interval between the date of the previous reference period, time t , and the reference date of the estimate, time $t + I$.
- $D_{(t,t+1)}^{a,s}$ = estimate of number of deaths of persons age a and sex s between time t and $t + I$;
- ${}^f D^{a,s}$ = number of deaths of persons age a and sex s during the corresponding quarter for which final death estimates f are available;
- ${}^f P^{a,s}$ = population estimate of persons age a and sex s at the beginning of the corresponding quarter for which final death estimates f are available;
- $P_t^{a,s}$ = population estimate of persons age a and sex s at the beginning of the quarter of estimation (time t).

The estimate of total number of deaths is then derived as follows:

Equation 3.4:
$$D_{(t,t+1)} = \sum_s \sum_a D_{(t,t+1)}^{a,s}$$

Monthly estimates are derived by re-distributing the given quarterly estimates using the most current final monthly distribution of births (or deaths).

The preliminary estimates are finalized when the vital statistics are available from Health Statistics Division.

3.2.4 Special treatment for preliminary estimates for Quebec and British Columbia

Quebec and British Columbia provide their most recent estimates of births and deaths to Statistics Canada. Their figures are used to produce preliminary estimates in Statistics Canada's publication. For the final estimates, the two provinces' births and deaths figures are taken from the vital statistics compiled by Health Statistics Division.

3.3 The use of birth and death estimates in other population estimates

Information from the Canadian Vital Statistics databases on births and deaths is used to produce other types of population estimates. For example, birth and death data (in addition to the other components of population change) are used to produce the population estimates by age, sex and marital status, which will be discussed in Chapter 9.

Births and deaths are also used to obtain census family estimates. Births may contribute to the formation of a family, whereas deaths may contribute to family dissolution. For a description of how birth and death estimates are used in the production of census family estimates, see Chapter 10.

Chapter 4

Immigration

This chapter provides the information on the data sources regarding immigration, and the methods used to produce estimates of immigrants by age, sex and marital status by province and territory. Information on the other four components of international migration can be found in the subsequent chapters.

4.1 Data sources and relevant concepts

The immigrant population refers to people who are landed immigrants in Canada. A landed immigrant is a person who is not a Canadian citizen by birth, but has been granted the right by immigration authorities to live in Canada on a permanent basis.

In Canada, immigration is regulated by the *Immigration and Refugee Protection Act* (IRPA) of 2002. This statute superseded the *Immigration Act*, which was passed in 1976 and amended more than 30 times in the years thereafter. Under the *IRPA*, there are three basic categories of permanent residents¹²: the economic class, the family class, and the protected persons category (or refugees).

Citizenship and Immigration Canada (CIC) collects and processes immigrants' administrative files. It then provides Statistics Canada with information from Field Operational Support System (FOSS) files. The information is used to estimate the number and characteristics of people granted permanent resident status by the federal government on a given date. In addition to basic demographic information about the age, sex, marital status and intended province of destination of immigrants, information is also available on the occupation, country of birth, and country of last permanent residence, mother tongue, level of education, official language abilities, etc.

Statistics on immigrants are usually tallied on or after the date on which they are granted permanent resident status or the right to live in Canada. For Demography Division, the terms immigrant and permanent resident are equivalent.

Table 4.1 shows the annual inflows of immigrants to Canada and by provinces and territories. Flows of new arrivals are volatile due to changes in Canada's immigration policy over the years and the supply of potential immigrants from any given country. Since the turn of this century, the range of new immigrants was around 220,000 to 260,000 annually. Three provinces have long attracted the vast majority of immigrants: Ontario, Quebec and British Columbia. For about twenty years, more

12. Children born abroad to Canadian parents who are out of the country are, by definition, Canadian citizens, and therefore are not included in estimates of immigration. Included however, are those persons who change status from non-permanent residents (i.e., permit/authorization holders or refugee status claimants) to landed immigrant status from within Canada. Although their migrations do not involve crossing Canadian borders, they are counted as non-permanent residents, upon their initial entry to Canada.

than 50% of the annual immigrants chose Ontario as their province of destination. Quebec and British Columbia each received about 17% of the newly arrived immigrants annually.

Table 4.1
Annual number of immigrants, Canada, provinces and territories, 1971/1972 to 2005/2006

Period (July 1 to June 30)	Canada	N.L.	P.E.I.	N.S.	N.B.	Que.	Ont.	Man.	Sask.	Alta.	B.C.	Y.T.	N.W.T.	Nvt.
1971/1972	117,036	729	174	1,709	1,081	17,690	62,258	5,086	1,479	8,243	18,402	70	115	...
1972/1973	138,526	751	171	1,923	1,428	19,861	75,267	5,352	1,555	9,300	22,570	124	224	...
1973/1974	217,456	1,128	377	3,064	2,279	33,213	120,369	7,769	2,194	13,789	32,986	97	191	...
1974/1975	209,283	995	284	2,342	1,998	31,339	112,876	7,334	2,514	15,688	33,618	97	198	...
1975/1976	170,028	1,013	222	2,004	2,241	28,363	85,648	6,489	2,604	15,905	25,261	94	184	...
1976/1977	130,931	648	200	1,854	1,425	25,034	63,267	5,154	2,380	13,819	16,961	46	143	...
1977/1978	100,967	466	174	1,219	751	16,169	49,986	4,426	1,906	11,634	14,048	63	125	...
1978/1979	84,518	377	168	967	689	14,833	40,768	3,410	1,673	8,942	12,537	51	103	...
1979/1980	143,825	681	300	1,885	1,530	24,385	64,536	7,207	3,722	17,195	22,144	93	147	...
1980/1981	127,238	483	146	1,271	963	18,787	54,648	6,327	2,961	18,567	22,879	108	98	...
1981/1982	135,339	424	148	1,470	883	24,052	57,870	5,165	2,371	20,642	22,126	93	95	...
1982/1983	101,404	356	139	937	662	18,700	44,951	4,515	1,913	13,668	15,424	60	79	...
1983/1984	88,592	311	103	939	558	15,268	40,252	3,958	1,861	10,749	14,470	58	65	...
1984/1985	83,925	300	116	1,049	596	13,568	40,442	3,478	2,091	9,843	12,320	47	75	...
1985/1986	88,657	298	129	974	625	15,944	43,158	3,906	1,918	9,211	12,390	43	61	...
1986/1987	130,880	355	165	1,170	678	25,452	70,252	4,078	2,131	10,637	15,823	70	69	...
1987/1988	152,211	434	165	1,212	580	24,620	85,125	4,836	1,990	12,661	20,455	65	68	...
1988/1989	177,632	431	139	1,452	732	29,493	98,070	5,294	2,228	15,024	24,584	84	101	...
1989/1990	203,357	483	181	1,454	954	37,763	108,863	6,766	2,201	17,992	26,496	116	88	...
1990/1991	221,382	614	149	1,542	738	45,791	115,213	6,342	2,284	17,903	30,642	64	100	...
1991/1992	244,281	704	165	1,927	804	51,600	129,706	4,792	2,538	17,001	34,808	118	108	10
1992/1993	266,890	806	161	2,599	748	48,357	145,962	5,419	2,563	18,975	41,014	128	121	37
1993/1994	235,360	704	139	3,084	589	35,964	120,181	4,525	2,283	18,207	49,432	108	119	25
1994/1995	220,738	615	200	3,726	676	26,768	119,730	3,811	2,191	16,515	46,260	100	127	19
1995/1996	217,478	557	127	3,397	646	29,490	115,961	3,666	1,825	13,870	47,728	89	114	8
1996/1997	224,857	479	185	3,111	673	27,991	119,401	4,030	1,775	13,783	53,235	97	79	18
1997/1998	194,459	411	123	2,590	717	27,242	106,419	3,093	1,599	11,648	40,456	78	68	15
1998/1999	173,194	368	125	1,624	752	27,739	91,899	3,293	1,752	11,243	34,274	67	46	12
1999/2000	205,710	425	142	1,674	609	30,250	116,744	4,207	1,671	12,865	36,946	79	86	12
2000/2001	252,533	453	190	1,761	883	36,705	150,040	4,834	1,848	16,200	39,489	48	72	10
2001/2002	256,334	422	145	1,615	768	39,079	152,836	4,825	1,821	16,658	37,985	67	94	19
2002/2003	199,193	314	89	1,255	648	34,926	109,827	4,906	1,558	13,621	31,931	54	59	5
2003/2004	239,082	536	267	1,707	760	44,552	127,945	7,417	1,894	17,111	36,719	58	105	11
2004/2005	244,579	541	312	1,708	865	43,417	129,790	7,686	2,097	17,442	40,575	58	79	9
2005/2006	254,359	450	343	2,199	1,387	41,983	133,116	8,884	2,112	19,869	43,858	76	73	9

... not applicable

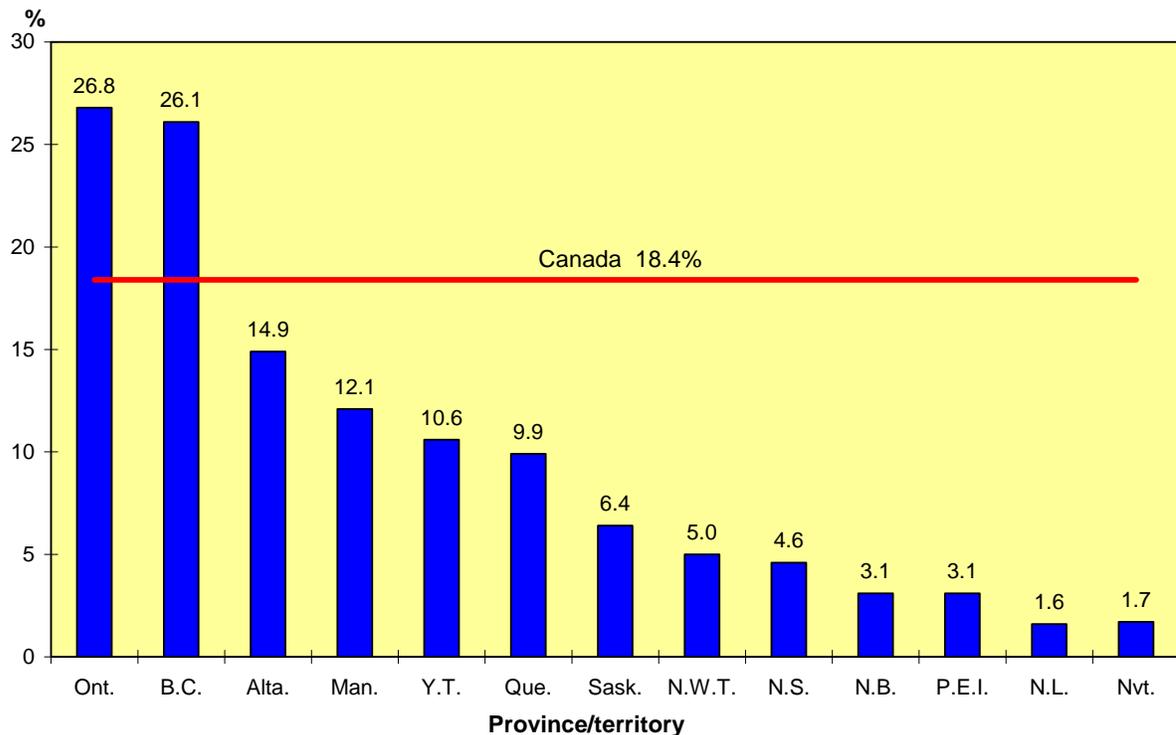
Note: The numbers of immigrants are final up to 2003/2004, updated for 2004/2005, and preliminary for 2005/2006.

Source: Statistics Canada, Demography Division.

Over time, there is an accumulated stock of immigrants living in Canada resulting from the flows on new arrivals, departures and returns of immigrants who subsequently left Canada. The census data are a source for information on stock of immigrants by place of residence, year of immigration, and age at immigration.

Figure 4.1 shows the provincial and territorial distribution of the population of immigrants from the 2001 Census. In 2001, immigrant population accounted for 18.4% of the total population of Canada. Ontario and British Columbia had a proportion of the immigrant population much above the Canadian average, 26.8% and 26.1%, respectively. This provincial and territorial distribution is different from the intended province of destination as indicated by immigrants on their records upon landing in Canada as the eventual place of residence of immigrants staying in Canada does not always correspond to their reported intended destination upon landing.

Figure 4.1
Immigrants as a percentage of total provincial and territorial population, 2001



Source: 2001 Census of Canada

4.2 Estimates of immigration, Canada, provinces and territories

Maintaining Canadian immigration statistics is statutory. Measuring the number of immigrants entering Canada in a given period (inflows) is straightforward, and adjustments to the data are not required. Information is available for each person entering Canada under landed immigrant status from CIC's administrative file.

Each month, CIC makes available to Statistics Canada, a data file containing the records of landed immigrants for the previous month, as well as any additions or updates to data already received. Given that there are typically few changes to the CIC data, the differences between preliminary and final estimates are very small.

For provincial and territorial level estimates, the file obtained from CIC identifies the province or territory of intended destination upon arrival, rather than the province or territory in which the immigrant actually settles. In a small number of cases, information on the province of destination is lacking. For these cases, the province of destination is distributed proportionately between the provinces and territories according to the distribution observed from immigrants for whom the information is available.

4.2.1 Immigration estimates by age, sex and marital status

The distribution of immigrants by age, sex and marital status is also straightforward, as these variables are available from the CIC file. The distribution only requires basic tabulation by age, sex and marital status. In the event of missing information, these cases are prorated according to the distribution for immigrants for whom the information is available.

4.2.2 Levels of estimates

The difference between preliminary and final postcensal estimates lies in the timeliness of the source used to estimate this component. Since the FOSS file is continually being updated, new calculations are carried out each year to update the immigration estimates. Immigration estimates are preliminary in the first year and updated the following year. They become final two years after the reference year.

Chapter 5

Non-permanent residents

Non-permanent residents (NPRs) are persons who have been legally granted the right to live in Canada on a temporary basis under the authority of a temporary resident permit, along with members of their family living with them. Non-permanent residents include foreign workers, foreign students, the humanitarian population and other temporary residents. The humanitarian population includes refugee claimants and temporary residents who are allowed to remain in Canada on humanitarian grounds and are not categorized as either foreign workers or foreign students. For Demography Division, the terms non-permanent resident and temporary resident are equivalent. Children born in Canada to parents of non-permanent resident status are considered as Canadians by birth and have all rights and privileges associated with citizenship.

Immigration data indicate that NPRs have grown considerably in the 1980s. In addition, they are staying longer in Canada. Because of their expanding role in Canada's society and economy, and in accordance with international practices¹³, NPRs were included in the target population for the first time in the 1991 Census¹⁴ (McKie, 1994). Though NPRs represent a small proportion of the Canadian population, about 1.3% of the total population, July 1, 2006, this is a heterogeneous group and they tend to congregate in Canada's large urban areas. Information on NPRs is useful for the review of employment policies and programs, as well as to plan education, health and other services. Consequently, the NPRs are a component of the population change due to international migration. This component is used with all the other international migration components to produce estimates of the international migration portion of the population of Canada and the provinces and territories¹⁵.

This chapter provides the information on the data sources on NPRs, and the methods used to produce estimates of NPRs by age, sex and marital status by province and territory. The following information mainly refers to the period after 2001.

5.1 Data sources and relevant concepts

The data required to produce the NPR estimates are obtained from Citizenship and Immigration Canada's (CIC) Field Operations Support System (FOSS) files. They include data on visitor permits, work authorizations, student authorizations, special temporary residents' permits, refugee status

13. The United Nations recommends that long-term residents (persons living in a country for one year or longer) be enumerated in the census.

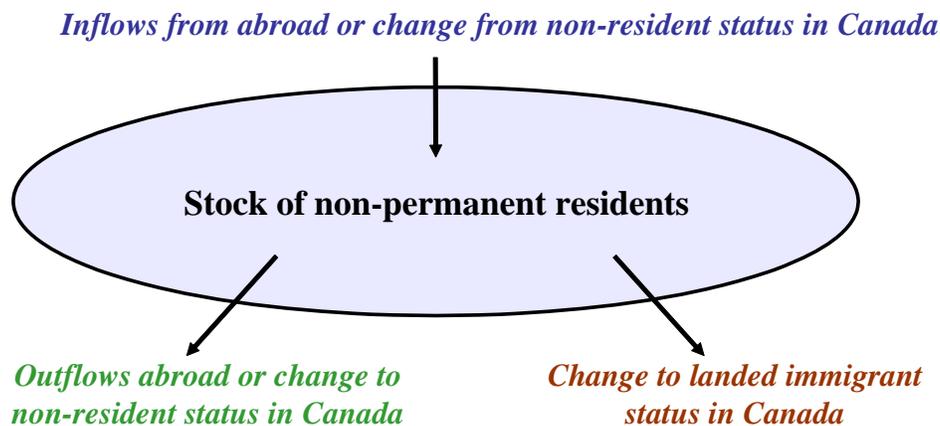
14. The only exception was the 1941 Census, in which Statistics Canada also included non-permanent residents in this population counts to account for persons entering Canada during World War II.

15. The United Nations has emphasized the importance of international migration and the maintenance of statistics on this topic. Thus, the inclusion of NPRs in Canadian population statistics also responds, in part, to UN recommendations which aim at establishing a universal definition of international migration for the purpose of internationally comparable statistics.

claims, landings¹⁶, deportations, applications for landing and proof of identity information for persons making refugee status claims.

Movements into and out of the NPR population are referred to as “flows”, as are status changes from permit/authorization holders to refugee status claimants or from permit/authorization holders or refugee status claimants to landed immigrant status¹⁷. “Stocks” represent the number of NPRs residing in Canada at any given time. As illustrated in Figure 5.1, from the most general perspective, persons can enter into the current stock of NPRs (i.e., “inflows”) either from abroad or from non-resident status within Canada (e.g. visitors). When a person leaves the NPR population (i.e., “outflows”), it is assumed that they have either left the country, become a non-resident, or become a landed immigrant.

Figure 5.1
NPR flows by broad type



In the population estimation program, NPR stocks are shown as a component of the total population estimates for a given date, while “net change” in NPR stocks over a certain period of time is included as a component of population change. “Net change” in NPR represents the variation in the number of NPRs between the two dates.

NPRs may be in Canada under any one of the following situations:

- (i) as holders of only permits or authorizations (they may concurrently hold more than one type);
- (ii) as persons who have only ever lodged refugee status claims; and
- (iii) as refugee status claimants who also possess one or more of the types of permits or authorizations that would qualify them for NPR status.

16. “Landings” refers to lawful permission to establish permanent residence in Canada. (*Immigration Act*, Chapter 2 – Interpretation), <http://laws.justice.gc.ca/>

17. Persons leaving the NPR population by obtaining landed immigrant status change the size of the NPR population, but have no impact on the size of the total population of Canada, as these persons become a part of the permanent resident population.

In cases where refugee status claimants concurrently hold some type of permit or authorization (e.g., a refugee status claimant is granted an employment authorization to help support themselves and their dependants within Canada), the refugee status supersedes the other NPR status. Consequently, references to “permit/authorization holders” excludes persons who have also made refugee status claims. The two major NPR sub-groups, permit/authorization holders and refugee status claimants, are used for estimation purposes. Dependants who were born abroad to members of these two sub-groups of NPRs are also included; however children born in Canada to NPRs are not, this population being covered by the birth component (see Chapter 3).

All persons in possession of a permit/authorization or claiming refugee status are assigned a Client Identification Number (CID) by Citizenship and Immigration Canada (CIC). This identification number is a key variable as it is unique for each person holding a permit/authorization or claiming refugee status and is maintained for every document issued by CIC. It is used in the production of estimates of NPRs.

Anyone who received non-permanent resident status prior to the reference date is counted in the NPR population. For refugee claimants, the date of their application is used as the date they receive NPR status. For permit/authorization holders the effective date is typically the start date of their permit. Permit holders and refugee claimants are excluded from the population if their permit has expired, if they receive permanent resident status, or if they are deported. In addition, refugee claimants are excluded if their file has been inactive for two years.

5.2 Estimates of stocks of non-permanent residents

The two major subgroups of the NPR population (permit/authorization holders and refugee status claimants) are administratively different; therefore, their estimates must be produced separately. The methods used in the production of estimates for permit/authorization holders are discussed first, followed by those used for refugee status claimants.

5.2.1 Stocks of permit/authorization holders

Permit/authorization holders (PHs) must either have had an official document signed in advance or have it signed upon entry to Canada, allowing them to reside in Canada on a temporary basis. Therefore, a person is considered part of this population if they possess a valid employment, student or Ministerial document on the reference date for the estimates. This means that the document must have been in effect prior to the reference date and valid until or past the reference date. If a person possesses more than one valid permit/authorization, the information from the permit/authorization for which the effective date¹⁸ is closest to the reference date is used.

Since dependants of permit/authorization holders are not required to obtain their own permits (though some do), the number of permit/authorization holders and their dependants is equal to the total number of persons covered by the documents.

18. The effective date is the date (or an approximation of) when the PH entered Canada as an NPR, or the date as of which a permit or authorization is extended.

Status as a permit/authorization holder terminates upon the expiry of the valid document, or when a PH is granted landing status (i.e., becoming a permanent resident), deported, or claims refugee status.¹⁹ Though the latter case affects the populations of the sub-groups of NPRs, the total NPR population is unaffected by this type of change in status.

The province of residence for each PH is obtained from the valid document. If a person has more than one valid permit/authorization, then the province of residence is taken from the permit/authorization for which the effective date is closest to the reference date. In some circumstances, the province of residence is not indicated on the document. In these cases, the province of the Canada Immigration Centre where the permit/authorization was issued is used. There are a small number of cases for which there is neither a province of residence nor a Canada Immigration Centre indicated. These cases are prorated according to the provincial distribution of PHs for which the province of residence is known or has been derived. In all cases, the province of residence for the principal PH is assumed for the dependants.

For each province/territory the stocks of PHs is a simple accounting of all valid permit holders and their dependants and can be expressed as follows:

PH = sum of all PHs and their dependants covered by the valid document on the reference date, as defined above

5.2.2 Stocks of refugee status claimants

A person is considered part of the refugee status claimant (RSC) population once a claim has been filed at a Canadian Immigration Centre. Effective with the proclamation of Bill C-86 on February 1, 1993, each person claiming refugee status is treated as a separate “case” (i.e., all claimants, including dependants are considered of separate identity and each is represented by a separate case, identified by a unique CID). Under the previous legislation (Bill C-55), it was not necessary for dependants to submit independent claims; therefore the number of RSCs was equal to the total number of persons represented by the claim of principal claimants.

Those persons who held legal temporary status in Canada as PHs prior to making a refugee status claim are included in the PH population until the date of their refugee status claim, at which time they are considered to be RSCs.

Since refugee status claims do not have an expiry date, a withdrawal of a claim, a deportation or being granted landing status has to occur for a claimant to leave the NPR population. For this reason, some assumptions are necessary to ascertain if a claim is still active.

19. Occasions arise where there appears to be a short interruption in a PH’s temporary stay in Canada. That is to say, a permit/authorization expires and there is a brief time lag between the expiry date and the effective date of the next permit/authorization. Since this is likely due to administrative delays in the issuance of permits/authorizations and extensions, interruptions of less than 31 days are disregarded and these persons are considered as having been continuously residing in Canada.

For those RSCs having applied for landed immigrant status:

- (i) it is assumed that they will leave the NPR population 2 years after their application for landing date unless they do not possess acceptable proof of identity (see (ii) below). This assumption is based on the fact that these applications typically take no longer than 2 years to process;
- (ii) it is assumed that if they do not have valid proof of identity²⁰, they will leave the NPR population 5 years after the date of application for landing. This assumption is in accordance with the recommendation of the Minister of Immigration.

For those RSCs who have not applied for landed immigrant status:

- (iii) it is assumed that they will leave the NPR population 2 years after their last communication with CIC. Records of every communication with each client are kept, whether it be a request for information on their status, to file a new application, to make an administrative change (e.g., change of address), etc. If there has been no activity over a 2 year period for an RSC, it is assumed that they are no longer an NPR. The assumption is that they either have left the country, or are deceased.

The province of residence for each RSC may be obtained from the claim record. If it is missing or invalid, if it exists, then information from permit/authorizations is used as a proxy (see previous section). If there is no information on the province of residence from these sources, then as with PHs, the province of the Canada Immigration Centre where the claim was made is used. If this is not available, these remaining cases are prorated according to the provincial distribution for RSCs where the province of residence is available (or derived).

Given the above, stocks of RSCs can be expressed as follows:

For each province/territory:

- (a) for claims submitted under Bill C-55 (between January 1, 1989 and January 31, 1993), the total number of persons covered by the active claim on the reference date; plus
- (b) for those submitted under Bill C-86 (effective February 1, 1993), the number of active individual claims on the reference date.

5.2.3 Level of estimates

The difference between preliminary and final estimates lies in the timeliness of the source used to estimate this component. Since the FOSS file is continually being updated, the figures are recalculated for each new release period to update the estimates of the net number of NPRs. NPR estimates are

20. Cases of persons arriving in Canada without proof of identity may arise in situations where they have left a country involved in civil unrest or war, with little or no belongings, including identification. Depending on the circumstances in their last country of residence, they may not be capable of obtaining appropriate identification even after they have entered Canada.

preliminary the first year and updated the following year. They are deemed final two years after the reference year.

5.3 NPR estimates by age, sex and marital status

5.3.1 Permit/authorization holders

Though permits and authorizations include information on the number of persons the document covers, data related to the age, sex and marital status are only available for the principal PH. Therefore, the information for principal applicants is obtained directly from the CIC files, while estimation is required for characteristics of their dependants.

Since data are available for all RSCs who filed under Bill C-86 (i.e., one application equals one person), the age, sex and marital status distributions of dependants of principal claimants who also hold permits, are used for the dependants of PHs.

5.3.2 Refugee status claimants

Since data are available for all RSCs who filed under Bill C-86, the data on sex, age and marital status for these claimants are obtained directly from their valid claim.

5.4 Net non-permanent residents as a component of change of total population estimates

The net change in NPRs from one period to another is determined by subtracting the NPR stock of the previous reference date from the current NPR stock.

Equation 5.1:
$$\Delta\text{NPR}_{(t,t+1)} = \text{NPR}_{(t+1)} - \text{NPR}_t$$

where

$(t, t+1)$ = interval between the date of the previous reference period, time t , and the reference date of the estimate, time $t+1$.

$\Delta\text{NPR}_{(t,t+1)}$ = change in NPR stocks between time t and time $t+1$;

$\text{NPR}_{(t+1)}$ = NPR stocks at time $t+1$;

NPR_t = NPR stocks at time t .

In the population estimation program, net change in NPRs is one of the components added to the base population to produce population estimates at a given date.

Chapter 6

Emigrants, net temporary emigrants and returning emigrants

Emigration results in a loss of population due to citizens or landed immigrants who leave the country to establish a residence in another country. This residence abroad may be intended as permanent or temporary. Relating to the concept of emigration, the calculation of Canada's population involves the estimation of emigrants, net temporary emigrants²¹ and returning emigrants.

Unlike immigration, there is no legal provision in Canada to maintain records for persons leaving the country either on a temporary or permanent basis. Therefore, estimates of the number of emigrants and persons living temporarily abroad and their characteristics must be derived through secondary sources such as Canadian administrative files or immigration statistics of other countries.

As emigration is the most difficult component to estimate, the methods of estimation are constantly evolving in an attempt to produce more accurate emigration estimates, given available sources of information and feasible methods. This chapter presents current methods used to produce the estimates of the three components of population change resulting from emigration.

6.1 Data sources and relevant concepts

Despite recommendations by the United Nations to establish a universal definition of international migrants, as well as a uniform method of recording information regarding emigration and immigration, not much progress has been made (United Nations, 1998). This has consequences on the ability to compare migration statistics between different countries, as well as the potential to use other nations' international migration statistics.

For the purpose of estimating Canada's population, the following three components relating to emigration are described. Emigration estimates require a distinction between those persons establishing a permanent residence in another country (i.e., emigrants), those persons living temporarily abroad, and finally the portion of emigrants who have returned to Canada. Different data sources and methods are necessary for the two types of emigrants, as well as for returning emigrants. Since 2004, estimates of emigrants, net temporary emigrants and returning emigrants are provided as separate components in publications on Canadian population estimates.

Emigrants are Canadian citizens or landed immigrants who have left Canada to establish a permanent residence in another country (sometimes referred to as "permanent emigration"). In the census, permanent emigration involves severing residential ties with Canada and acquiring permanent residency in another country. Because of Demography Division's use of administrative and financial data, this will also include severing financial and social ties with Canada. For example, persons

21. Before 2004, net temporary emigrants were referred as net change in the number of persons living temporarily abroad in the Statistics Canada population estimation program.

traveling to the United States may be considered permanent emigrants if they acquire permanent resident status there, but are considered temporary emigrants if they still hold a visa or are on visitor status. For estimation purposes, emigration also includes those who inform Canada Revenue Agency with their intentions regarding social programs (e.g., Canada Child Tax Benefit) and income tax. Permanent emigration is a misnomer as it is not necessarily irreversible; emigrants can always decide to return to Canada.

Emigrants are estimated from administrative sources in terms of the “gross flow” of migrants out of Canada. The U.S. Department of Homeland Security²², Office of Immigration Statistics provides data on Canadians who acquire landed immigrant status in the U.S. This data source is used in estimating emigration to the United States. In order to estimate emigration to other countries, information on notification of departure from the Canada Child Tax Benefit program (CCTB) and tax data from Canada Revenue Agency (CRA) is used.

Canada Child Tax Benefit data (CCTB), combined with personal income tax data, cover emigration better than tax data alone. The CCTB program identifies emigrants through their tax return and/or from their notification of departure to CRA. The notification of departure can cover, in part, families that will not necessarily file a tax return for the year of their departure. The combination of CCTB and tax data allows for an estimate of child and adult emigration and, using Demography population estimates, provides an estimate of the non-covered CCTB child population. Tax data on filers and their dependants are provided by Small Area and Administrative Data Division (SAADD) of Statistics Canada, based on files received from CRA.

Some people leave Canada to live temporarily in another country while not maintaining a usual place of residence in Canada. Others who were temporarily outside Canada return. The net result of those departures and returns is the component known as “**net temporary emigration**”. Data from the Reverse Record Check (RRC), the most important census coverage study, are used to estimate the number of persons leaving the country temporarily; while data from the census, combined with the returning emigrant estimates, are used to estimate the number of temporary emigrants returning.

Temporary migration constitutes many movements in and out of the country. This would not concern the overall population estimates if the net effect was nil or negligible. However, the census coverage study (Reverse Record Check) in 1996 has shown that this is not the case. The decision to account for persons living temporarily abroad was based on evaluations of the 1996 error of closure for postcensal population estimates and its components against Reverse Record Check estimates. These evaluations concluded that omitting departures of Canadians for temporary residence abroad and their consecutive returns to Canada has an important impact on the quality of the estimate of the country’s population (Michalowski, 1999). Including net change in temporary emigration in the population estimation program is a practice that started in 1998, with revisions to estimates back to 1996.

Returning emigrants are Canadian citizens or landed immigrants having previously emigrated from Canada and subsequently returned to Canada to re-establish a permanent residence. Again, data from the CCTB program are used in estimating returning emigrants. The rules governing returning emigrants may in fact be different from Canadian born returning emigrants and those who previously

22. The U.S. Department of Homeland Security, Office of Immigration Statistics is formerly called Immigration and Naturalization Services (INS) of the U.S. Department of Justice.

had received landed immigrant status. But as of yet, we do not differentiate between the two populations.

6.2 Estimates of emigrants, net temporary emigrants and returning emigrants, Canada, provinces and territories

6.2.1 Final estimates

A. Emigrants

Estimates of emigrants refer to permanent emigrants only. As mentioned earlier, emigrants are counted in terms of flows, i.e., each movement in and out of the emigrant population are counted.

Statistics from the U.S. Department of Homeland Security, Office of Immigration Statistics, administrative data from the Canada Child Tax Benefit program (CCTB) and the personal income tax data from the Canada Revenue Agency (CRA) are used to produce estimates of permanent emigrants. As the CCTB program is not universal and does not provide direct information on the number of adult emigrants, four adjustment factors are used to compensate for the population it does not measure:

1. the delay in the completeness of the files on emigrant CCTB-children;
2. the partial coverage of the program, as the data do not include persons who do not apply to the CCTB program or are not eligible²³;
3. the differential propensity to emigrate between children covered by the program compared to those who are not; and
4. the number of adults emigrating, based on past ratios of the rate of child emigration to adult emigration.

Each of these adjustment factors makes use of CCTB and tax data from CRA, and the estimation methods for child and adult emigrants are described in the next section.

Emigrant children

The CCTB file provides numbers of dependant children (under 18 years of age) whose parents, who are eligible and have applied for the CCTB, have become non-residents as defined by CRA. These data are available on a monthly basis but provided by CRA to Statistics Canada annually for each province and territory. As mentioned, the CCTB program is not universal; therefore the data are incomplete in terms of estimation of total child emigrants and require adjustments.

The first correction factor is to take into account the incompleteness of the emigration information from the CRA's CCTB data files. Based on comparisons done with files that were two, three and four

23. Canadian citizens, permanent residents and non-permanent residents who have been in Canada for a year and subject to Canadian taxation are eligible to CCTB.

year after the reference period, it appeared that the CRA files could be considered complete about four years after the reference period. Therefore, an adjustment is made to the data when they are finalized only two years after the reference period. The same adjustment factor is applied to the monthly data of each province/territory.

The second correction factor, that is the adjustment for partial coverage, corrects the number of children who would be included in the CCTB program if it was universal. The adjustment starts with a correction factor that is applied to the population registered for the CCTB program to account for its shortfall in total coverage: the ratio of the number of children registered for the CCTB program to the number of children in the total population, as estimated by Demography Division for each month and province and territory²⁴.

The third correction factor is used to adjust the differential “emigration propensity” between children registered for the CCTB program and that of all children for each province and territory, on an annual basis. This factor is calculated based on tax data provided by SAADD, which identify emigrant families by comparing addresses on income tax returns for two consecutive years and include information on persons registered for the CCTB program during the taxation year. It was found that the estimated emigration rate of CCTB registered families was consistently lower than the rate for all families with children. Consequently, an adjustment that reflects the higher propensity of emigration for non-CCTB children is done to compensate for the error. A moving average measured over a three-year period is used in computing this factor for each province. To eliminate variations due to small numbers in each of the Atlantic province, the differential propensity factor is estimated for these provinces as a group. In addition, the estimated factor for Canada is used for the territories.

The formula to estimate child emigrants by province/territory is as follows:

For each province/territory:

Equation 6.1:
$${}_{j}^{CCTB}E^{0-17} = {}_{j}CCTB_EM^{0-17} \times \frac{1}{{}_{j}^{CCTB}\mathfrak{R}} \times {}_{j}G$$

where

- ${}_{j}^{CCTB}E^{0-17}$ = emigrant children aged 0 to 17 from province/territory j based on CCTB data;
- ${}_{j}CCTB_EM^{0-17}$ = number of CCTB/Tax children moving from province/territory j;
- ${}_{j}^{CCTB}\mathfrak{R}$ = coverage rate of CCTB program for province/territory j;
- ${}_{j}G$ = adjustment factor for emigration propensity of non CCTB/Tax children for all provinces except the Atlantic provinces where the factor is measured as a whole and in the Territories where the Canadian factor is used;
- CCTB = CCTB data.

24. Based on annual CCTB and tax data, estimates of coverage rate of CCTB program by month can be derived.

The equation above includes the adjustments for incomplete coverage of CCTB-registered children (Equation 6.2) and for the differences in the propensities of CCTB-registered and non-registered subpopulations to emigrate (Equation 6.3).

The coverage rates are calculated on an annual basis as follows:

For each province and territory:

Equation 6.2:
$${}_{\text{CCTB}}\mathfrak{R}_j = \left[\frac{{}_{\text{CCTB_Tax}}\mathbf{P}_j^{0-17}}{{}_j\mathbf{P}^{0-17}} \right]$$

where

- ${}_{\text{CCTB}}\mathfrak{R}_j$ = coverage rate of CCTB program for each province/territory j;
- ${}_{\text{CCTB_Tax}}\mathbf{P}_j^{0-17}$ = number of children aged 0 to 17 years registered for the CCTB program in each province/territory j, as modeled from tax data;
- ${}_j\mathbf{P}^{0-17}$ = population estimated by Demography Division of children aged 0 to 17 years in each province/territory j.

The differential propensity to emigrate is obtained by dividing the emigration rates for all children by the emigration rates for children registered for the CCTB program, as follows:

For each province and territory:

Equation 6.3:
$${}_j\mathbf{G} = \frac{{}_j\text{RateofTax_EM}^{0-17}}{{}_j\text{RateofCCTB_Tax_EM}^{0-17}}$$

$${}_j\mathbf{G} = \frac{\left[\frac{{}_{\text{Tax}}\mathbf{E}_j^{0-17}}{{}_{\text{Tax}}\mathbf{P}_j^{0-17}} \right]}{\left[\frac{{}_{\text{CCTB_Tax}}\mathbf{E}_j^{0-17}}{{}_{\text{CCTB_Tax}}\mathbf{P}_j^{0-17}} \right]}$$

where

- ${}_j\mathbf{G}$ = adjustment factor for emigration propensity of non-CCTB/Tax children for each province j of origin except the Atlantic provinces where the factor is measured as a whole and in the Territories where the Canadian factor is used;
- ${}_j\text{RateofTax_EM}^{0-17}$ = emigration rate of all children from each province j except the Atlantic provinces where the rate is measured as a whole and in the Territories where the Canadian rate is used;
- ${}_j\text{RateofCCTB_Tax_EM}^{0-17}$ = emigration rate of CCTB-registered children from each province j except the Atlantic provinces where the rate is measured as a whole

and in the Territories where the Canadian rate is used, as modeled from tax data;

- $Tax_j E^{0-17}$ = emigration of all children from province, region or Canada j, according to income tax data;
- $Tax_j P^{0-17}$ = SAADD estimate of children in province, region or Canada j based on income tax files;
- $CCTB_Tax_j E^{0-17}$ = emigration of CCTB-registered children from province, region or Canada j, as modeled from tax data;
- $CCTB_Tax_j P^{0-17}$ = CCTB-registered children in the population of the income tax file provided by SAADD, by province, region or Canada j.

Data on Canadian adults and children who immigrated to the United States are available from the Department of Homeland Security on an annual basis. The Homeland Security data provide quarterly flow of immigrants to the United States, including those emigrating from Canada. Since the Homeland Security data are of better quality than the estimates of emigration from the CCTB model, the CCTB estimates are replaced by the Homeland Security data for emigration to the United States. Combining the estimates of child emigrants from CCTB data and the Homeland Security data on children who moved to the United States, estimates of children emigrating to countries other than the United States can be derived as follows:

Equation 6.4: $OC E^{0-17} = CCTB E^{0-17} - HS E^{0-17}$

where

- $OC E^{0-17}$ = emigrant children to countries other than the United States;
- $CCTB E^{0-17}$ = emigrant children based on CCTB data;
- $HS E^{0-17}$ = emigrant children to the United States based on Homeland Security data.

Emigrant adults

Estimates for adults emigrating to the United States are taken directly from the Homeland Security data. As the CCTB program does not provide direct information on emigrating adults, hence an adjustment factor was used to estimate the number of adults emigrating to countries other than the United States, based on the emigration rate of children to other countries²⁵.

The formula for estimating the number of adult emigrants to a country other than the United States is:

25. This ratio was obtained through the historical analysis of adult and child emigration rates based on SAADD data and HS data.

$$\text{Equation 6.5: } {}^{\text{OC}}E^{18+} = \frac{{}^{\text{OC}}E^{0-17}}{P^{0-17}} \times P^{18+} \times \sum_t \frac{\text{RatioAC}_t}{3}$$

where

- ${}^{\text{OC}}E^{18+}$ = total emigrating adults to countries other than the United States;
 $\frac{{}^{\text{OC}}E^{0-17}}{P^{0-17}}$ = emigration rate of children moving to countries other than the United States;
 P^{18+} = population estimate of adults at the beginning of the period;
 RatioAC_t = ratio of adult emigration rate over child emigration rate at the reference date, time t;
t = the year index for the current reference date up to the two previous reference dates (t, t-1, t-2).

The RatioAC of adult migration rate over child migration rate is an average of ratios from two sources; tax and Homeland Security²⁶. It is calculated for any given year as follows:

$$\text{Equation 6.6: } \text{RatioAC} = \frac{({}^{\text{Tax}}\text{RatioAC} + {}^{\text{HS}}\text{RatioAC})}{2}$$

where

${}^{\text{Tax}}\text{RatioAC}$ is the adult emigration rate divided by the child emigration rate taken from tax data. It is given by:

$$\text{Equation 6.7: } {}^{\text{Tax}}\text{RatioAC} = \frac{\left[\frac{{}^{\text{Tax}}E^{18+}}{{}^{\text{Tax}}P^{18+}} \right]}{\left[\frac{{}^{\text{Tax}}E^{0-17}}{{}^{\text{Tax}}P^{0-17}} \right]}$$

where

- ${}^{\text{Tax}}E^{18+}$ = total emigrating adults measured from tax data;
 ${}^{\text{Tax}}P^{18+}$ = SAADD estimate of adults based on income tax files;
 ${}^{\text{Tax}}E^{0-17}$ = emigration of all children according to income tax data;
 ${}^{\text{Tax}}P^{0-17}$ = SAADD estimate of children based on income tax data;

and

26. Because of past legislative changes that have occurred in the US and other unforeseen events, sometimes the information was not available at production time. If either sources of the ratio of child emigration over adult emigration is missing or late for the production cycle of the estimates, then the past calculation of this ratio is used as a proxy.

$^{HS}RatioAC$ is the adult emigration rate divided by the child emigration rate taken from the Homeland Security data. It is given by:

$$Equation\ 6.8: \quad ^{HS}RatioAC = \frac{\left[\frac{^{HS}E^{18+}}{^{Dem}P^{18+}} \right]}{\left[\frac{^{HS}E^{0-17}}{^{Dem}P^{0-17}} \right]}$$

where

- $^{HS}E^{18+}$ = total emigrating adults measured from Homeland Security data;
- $^{Dem}P^{18+}$ = estimate of adults from Demography Division one year prior to the reference date;
- $^{HS}E^{0-17}$ = emigration of all children according to Homeland Security data;
- $^{Dem}P^{0-17}$ = estimate of children from Demography Division one year prior to the reference date;

Finally, the total number of adult emigrants equals:

$$Equation\ 6.9: \quad E^{18+} = ^{OC}E^{18+} + ^{HS}E^{18+}$$

where

- E^{18+} = total emigrating adults;
- $^{OC}E^{18+}$ = number of adult emigrants to countries other than the United States;
- $^{HS}E^{18+}$ = number of adult emigrants to the United States, according to Homeland Security data.

The process of combining CCTB and Homeland Security data produces slightly inconsistent data in that the definitions of emigration used by the two data sources are not entirely the same. The Homeland Security definition of an immigrant to the United States (i.e., emigrants from other countries) is dependant on a person acquiring legal permanent resident status in their country. On the other hand, the CCTB usage does not subscribe to a definition of permanent resident status in another country to be considered an emigrant: one simply loses their resident status in Canada for income tax purposes. It is assumed that the difference between the two definitions is not sufficiently large to produce significant bias in the estimates.

Since the Homeland Security data do not provide information about the province of origin of Canadian emigrants, provincial and territorial level estimates of adult emigration must be derived from an alternate source. To estimate the adult emigrant distribution by province or territory of origin, the provincial and territorial distribution for child emigrants is assumed to be the same for the entire population.

Estimates of emigrants by age, sex and marital status

The choice of data source for each distribution (age, sex, marital status) was based on comparisons between the various data sources for emigration statistics (annual SAADD data and Homeland

Security data, as well as the Reverse Record Check (RRC), which provides information on emigration every five years).

The national estimate of emigrants is first apportioned by sex. For emigrants to the United States, the breakdown is directly available from the Homeland Security data. For emigrants to other countries, the Homeland Security distribution is used as a proxy. This distribution is closer to what was measured with the RRC than is the SAADD estimates by sex.

Estimates of male and female emigrants are then apportioned by age groups, using SAADD's distribution for the following five broad age groups: 0 to 17, 18 to 24, 25 to 44, 45 to 64, and 65 years and over. The Homeland Security data are used to further distribute the data into five-year age groups. Finally, Sprague's multipliers are applied to split the estimates into single years of age²⁷. Since provincial and territorial level data are either unreliable or not available from these sources, the national age/sex distribution is used for all provinces and territories.

To further distribute the estimates by marital status, the Homeland Security distribution of emigrants to the United States is applied to all emigrants in each of the provinces and territories.

B. Net temporary emigration

Net temporary emigration is the net result of the movement of people leaving Canada to live temporarily in another country and of those people that have returned to Canada. The following steps are used to estimate their monthly numbers by province/territory:

1. Estimate of the number of departures at Canada level is derived from the Reverse Record Check (RRC), the most important census coverage study. The RRC provides an estimate of the number of people who left Canada temporarily²⁸ during an intercensal period and who are still out of the country at the end of the period.
2. Estimate of the number of temporary emigrants returning is done in two steps:
 - a. The number of all returning emigrants for Canada is taken from the census. The census provides the number of persons who resided outside Canada at the previous census and who have since returned to the country during the intercensal period;
 - b. From the estimate of all returning emigrants for Canada (census estimates Step 2.a) we subtract Demography Division's estimate of returning permanent emigrants. The resultant estimate is the number of temporary emigrants returning.
3. Estimate of net temporary emigration for Canada is derived by subtracting returning persons (Step 2.b) from departures (Step 1).
4. The estimate derived in Step 3 is then distributed by province/territory according to the provincial and territorial distribution of departures of temporary emigrants based on the RRC²⁹. The number for the Atlantic provinces is estimated as a group and redistributed proportionately

27. Sprague's multipliers are interpolation coefficients used to subdivide data. For a detailed description of the method of Sprague's multipliers, see Shryock et al (1976).

28. Emigration is defined as temporary based on the intention of returning and the time spent outside the country.

29. Special estimates are done for the territories as the RRC does not provide estimates.

to each province according to their respective population size. The same is done for the three territories.

- The provincial and territorial estimates are disaggregated equally into annual estimates for each of the five years of the intercensal period. The monthly estimates are assumed to have a seasonal distribution. This distribution is modeled after the seasonal patterns observed for permanent emigration. The seasonal patterns are assumed to be between an even flow of net temporary migrants and the proportional flow as measured for permanent emigration which is expressed as an average of the two seasonal patterns. The mathematical expression that gives each monthly flow is as follows:

Equation 6.10:

$$NTE_m = \frac{\left[\left(\frac{NTE}{12} \right) + \left(\frac{PE_m}{PE} \times NTE \right) \right]}{2}$$

where

- NTE_m = the number of net temporary emigrants for the month m ;
 NTE = the annual number of net temporary emigrants;
 PE_m = the number of permanent emigrants for the month m ;
 PE = the number of annual permanent emigrants.

Net temporary emigration can only be estimated for the intercensal period preceding the most recent census. Postcensal estimates of net temporary emigration is assumed to be the same as those estimated in the previous intercensal period for each province and territory as no other source of information is available. They remain unchanged until the completion of the RRC in the next census (i.e. approximately two years after the census).

Estimates of net temporary emigrants by age, sex and marital status

To obtain the estimates of net temporary emigrants by age, sex and marital status, the same distributions as observed for permanent emigration are used.³⁰

C. Returning emigrants

A returning emigrant is a person returning to Canada, after having been classified as an emigrant. The number of returning emigrants is estimated using CCTB data covering the dependant children of CCTB recipients who have returned to Canada. A citizen or permanent resident who has emigrated regains eligibility for the CCTB if he or she re-establishes residential, economic and social ties in Canada. A returning child emigrant is identified by the presence of both a departure date and return date on the CCTB file, as well as the parent's residency status. As with emigration, a person's return to Canada relies on their resident status for income tax purposes. It is assumed that most of the returnees to the CCTB program had emigrated before.

30. Although the Reverse Record Check provides age-sex data on temporary emigration, they are unreliable due to the large sampling variances at this level of detail.

CCTB data for returning emigrants include children born outside Canada. These children are included in the returning emigrant population because the majority of foreign births to Canadian parents are not included in vital statistics (according to provincial registry offices). Their inclusion makes up for what would otherwise be an underestimation of population estimates for the reason of incompleteness of data on births. As is the case with emigration, estimates of the number of returning emigrant children and the number of returning emigrant adults are calculated separately.

Returning emigrant children

The CCTB file provides for each province/territory, monthly numbers of dependent children (under the age of 18 years) of CCTB recipients who have returned to Canada after a period of emigration.

As with emigrant children, the CCTB data for returning child emigrants are adjusted with a factor reflecting the program's coverage. In this case, it is assumed that the ratio of CCTB-eligible to CCTB-ineligible returning emigrants is the same as the ratio for emigrants. Hence, the numbers for returning CCTB-eligible children are adjusted with the same differential propensity factors that are used for emigration. The formula provided earlier used to compute the estimate of child emigrants (Equation 6.1) is used to estimate returning child emigrants, where $({}^{\text{CCTB}}_j \text{EM}^{0-17})$ the number of child emigrants by province/territory, according to CCTB data, is replaced by $({}^{\text{CCTB}}_j \text{RE}^{0-17})$ the number of children returning to Canada.

Returning emigrant adults

The number of adults returning to Canada after emigrating can be indirectly estimated using SAADD estimates based on income tax files or census data. Both sources include returning permanent emigrants, returning temporary emigrants and persons who may have maintained residential ties with Canada. However, neither source can be used directly to provide the number of returning adult permanent emigrants. It is possible to obtain the adult/child ratio in the returning emigrant population as follows:

Equation 6.11:
$$\tau^{\text{RE}} = \frac{{}^{\text{C}} \text{RE}^{18+}}{{}^{\text{C}} \text{RE}^{0-17}}$$

where

- τ^{RE} = adult to child ratio of the returning emigrant population;
- ${}^{\text{C}} \text{RE}^{18+}$ = number of returning adult emigrants based on census counts;
- ${}^{\text{C}} \text{RE}^{0-17}$ = number of returning child emigrants based on census counts.

The number of returning adult emigrants is estimated by multiplying the estimate of returning child emigrants based on CCTB data by the adult/child ratio from Equation 6.11.

Equation 6.12:
$$\text{RE}^{18+} = {}^{\text{CCTB}} \text{RE}^{0-17} \times \tau^{\text{RE}}$$

where

- RE^{18+} = estimated number of returning adult emigrants;
 $CCTB RE^{0-17}$ = estimate of returning child emigrants based on CCTB data;
 τ^{RE} = adult/child ratio in the returning emigrant population based on census count.

The estimate of the number of returning children is produced monthly for each province and territory. The provincial and territorial distribution of returning children is assumed to hold for returning adults.

Estimates of returning emigrants by age, sex and marital status

The estimates of returning emigrants are distributed by age, sex and marital status using the most recent census data on Canadians who were living abroad one year prior to the census, excluding immigrants who arrived that year.³¹ The census distribution by single years of age and sex is applied to children aged 0 to 17 years. For adults 18 years and over, census distributions by age, sex and marital status is used.

6.2.2 Preliminary estimates

Final data for emigration estimates are not received until approximately two years after the reference period. As such two reference periods, $(t, t + 1)$ and $(t + 1, t + 2)$ elapse before final data are received for the first reference period $(t, t + 1)$.

For the period $(t, t + 1)$, preliminary estimates of child emigrants are produced using partial, but nearly final CCTB data for the reference period. For each province and territory, the CCTB coverage factor and the differential emigration propensity are calculated, as in Equation 6.1. The number of adults emigrating to the United States is assumed constant from the most recent final Homeland Security data. For adults emigrating to countries other than the United States, the same procedure is used as for final emigration estimates (see Equations 6.5 to 6.9).

For the period $(t + 1, t + 2)$, preliminary estimates of emigration are produced as described for the period $(t, t + 1)$, assuming constant CCTB numbers from $(t, t + 1)$ as well as those from Homeland Security. Though the same data from these two sources are used, the resulting estimates of emigration differ slightly, due to the change in the estimates of the population of children and adults over the two periods.

31. Returning Canadians can also be identified by census questions on place of residence five years earlier, once immigrants who arrived in Canada during this period have been excluded. It is judged that demographic characteristics of returning Canadians at the time of return are better reflected by the data obtained through information on their residence one year ago.

Chapter 7

Interprovincial migration

Interprovincial (or interterritorial) migration represents movements from one province or territory to another, involving a change in usual place of residence. Subprovincial migration (or intraprovincial migration) also involves a change in usual place of residence, but these movements occur between smaller units of geography such as census divisions (CDs) or census metropolitan areas (CMAs). The term “internal migration” may refer to either interprovincial or subprovincial migration, or both. This chapter focuses on the interprovincial migration component. For information on subprovincial migration, see Chapter 8 on Subprovincial Estimates of Population.

As is the case for emigration, there is no provision for recording interprovincial migration in Canada. Consequently, this component of population change has to be estimated using administrative data. The methods used to prepare estimates of interprovincial migration will be discussed in this chapter.

7.1 Data sources and relevant concepts

Canada Revenue Agency (CRA) provides Statistics Canada with data from the Canada Child Tax Benefit (CCTB) program and personal income tax returns, which are used to estimate interprovincial migration. Preliminary estimates are derived monthly from CCTB data, while final estimates are derived annually from T1 Family File (T1FF)³².

Since 1976, personal income tax records have become the official data source for final migration estimates. The population covered by tax data is more comprehensive than that of the CCTB; in general, interprovincial migration estimates based on tax data are considered to be of higher quality than those produced using data from the CCTB. However, tax data are not timely enough for use in preliminary estimates. Small Area and Administrative Data Division (SAADD) of Statistics Canada receives tax data only after annual processing, compared to CCTB data which are available on a monthly basis.

7.2 Estimates of interprovincial migration, Canada, provinces and territories

7.2.1 Final estimates of interprovincial migration

Final estimates of interprovincial migration are produced by SAADD, using personal income tax data. By comparing the place of residence at the time of filing, it is possible to identify for those

32. The T1 Family File (T1FF) is derived from CRA’s T1 file by Small Area and Administrative Data Division (SAADD) of Statistics Canada.

persons who file returns for two consecutive years, those who move, their place of origin and their destination. The development of migration data involves four main steps³³:

1. geo-coding of tax records;
2. estimation of non-filing dependants of tax filers, by age groups and gender;
3. identification of migrant tax filers by age group and gender;
4. adjustment for the population not covered by the Canada Revenue Agency Taxation system.

The four steps of the estimation methodology are described below.

Step 1 – Geo-coding

The migration data developed from the taxation records are estimates of migration flows between census divisions (CDs) or census metropolitan areas (CMAs). The geographic coding of CDs and CMAs on the tax records is done primarily on the basis of the postal code, which is a part of the mailing address. Tax records also contain a “locality code” assigned by Revenue Canada based on place name, which is a combination of Standard Geographic Classification codes used by the census. About 99% of all tax records were filed with a postal code, or were assigned one, based on the filer’s address by matching the address to a file similar to the Postal Code Directory from Canada Post. As the tax return is usually submitted several months after the end of the tax year, the postal codes correspond to those existing in the spring of the year following the tax year.

Step 2 – Estimation of dependants

Since the tax records source file has no direct information on the number and characteristics of non-filing dependants, this information must be imputed. Following the 1988 tax reforms, the estimation of tax filers’ dependants is obtained from the T1 Family File (T1FF). The family system creates families by linking all filing family members together, using spousal social insurance numbers, marital status, and matched addresses. This system imputes a non-filing spouse whenever a filer has declared himself/herself married but was not linked to a filing spouse. Children are imputed based on CCTB data, which carry the social insurance number of the parent receiving the benefit. Newborns are added to tax filer records from the Statistics Canada’s Vital Statistics birth database, and in addition a historical file of imputed children is created from the previous year’s tax file³⁴.

The family data from the tax file are then used to create a file of individual tax filers for the migration system. The migration file contains information for each tax filer and assigns dependants to each tax filer based on the number of filers in the family, and the number of dependants. If only one parent in the family has filed a tax return, then all dependants, including the spouse, if one exists, are added to the tax filer’s record. If both parents filed tax returns, and their family record contained dependant children, then each tax filer would receive half the count of dependants.

33. For a detailed description of the methodology, see Statistics Canada, 2005 and Reid, 1998.

34. The Historical Family File, updated annually by SAADD, consists of information of known family relationships obtained from tax returns.

If not already present, the age of the dependant spouse is estimated in the family system. The age of imputed children is provided (date of birth) on CCTB records, birth files from Vital Statistics, and from the Historical Family File. The family system does not impute gender for non-filing children. The gender for each dependant child is assigned a male/female ratio based on information from the most recent census.

Step 3 – Identifying migrant tax filers

After assigning the geography codes and non-filing dependants to each tax filer, records for two consecutive years are matched by social insurance number, and migrant tax filers are identified by comparing current and previous geography codes (CDs or CMAs). Once the “at risk” population is tabulated, only records with a change in geographic coding are kept further in the migration process. Thus, with the exception of immigrants, it is only possible to determine migrant status for those who file two consecutive years. The migrant tax filers are identified by comparing the place of residence at the time of filing, in two consecutive years. The identification of migrants is based on their address at the time of filing, and thus the migration period is not precisely one year, but it is assumed to be close. The assumed reference period is around April to April, of the year following each tax year. Tax filers’ non-filing dependants are assumed to have the same migration behaviour as that of the filers to whom they are assigned.

Step 4 – Coverage adjustment factor³⁵

The final step in the estimation process is an adjustment for coverage, done at the CD/CMA level. This is required to estimate those migrants who do not file two consecutive personal tax returns, which involves the calculation of coverage adjustment factor (or inflation factor). Population estimates by CD/CMA are used to create coverage ratios. For migration estimates up to 2000/2001, provincial adjustment ratios were used in place of the CD/CMA ratio in the few cases where coverage was abnormally high or low. Beginning with 2001/2002 migration data, high and low coverage cases were identified with a new methodology and a Canadian adjustment ratio was used in place of the CD/CMA ratio.

The adjustment ratios are applied to the counts of out-migrants derived in Step 3 to obtain an estimate of total migration. The basic assumption is that the population not covered by the taxation system has the same migration rate as that covered by it.

7.2.2 Preliminary estimates of interprovincial migration

Preliminary estimates of migration between provinces and territories are produced using CCTB data, along with data from personal income tax records, processed by SAADD. Because the CCTB program is not universal and does not provide direct information on the number of adult migrants, this estimation requires three adjustment factors to take into account:

35. For 2001/2002 and subsequent years, the adjustment method is slightly modified, for details, see Wilkinson, 2004.

1. the program’s partial coverage (i.e., persons who do not apply to the CCTB program or are not eligible);
2. the differential propensity to migrate between children who are receiving the CCTB and children who are not;
3. the differential propensity to migrate between adults and children.

Each of these adjustment factors makes use of data from CCTB and SAADD’s T1FF. The estimation methods for child and adult interprovincial migration rates are described in the next section.

Interprovincial migrant children

Since income tax returns are not available at the time preliminary estimates are produced, the estimation of preliminary interprovincial migration is based on CCTB administrative files, which provide counts of child migrants (aged 0 to 17) receiving benefits under the program. The data provide monthly information on the changes of usual residence involving provincial or territorial boundaries of CCTB-registered children and the total number of children registered for the CCTB program by province and territory. As mentioned, the CCTB program is not universal; therefore the data are incomplete in terms of estimation of total child migrants and require adjustments.

The estimation of interprovincial migrant children involves two adjustment factors. The first factor reflects the CCTB program’s coverage: the ratio of the number of children as estimated by Demography Division, to the number of children actually covered by the CCTB program, for each month and province and territory. The second factor estimates the differential propensity to migrate between children who are receiving the CCTB and children who are not. This factor is obtained by comparing the out-migration rates of children receiving the CCTB with the rates for all children (aged 0 to 17). This factor is calculated for each province and territory and is based on the two last available years of SAADD’s T1FF.³⁶

The formula to estimate children migrating between provinces/territories is as follows:

Equation 7.1:
$${}_{j,k}I^{0-17} = {}_{j,k}CCTB_IM^{0-17} \times \frac{1}{{}_{j}^{CCTB}\mathfrak{R}} \times {}_{j}G$$

where

- ${}_{j,k}I^{0-17}$ = number of migrant children from province/territory of origin j to destination province/territory k;
- ${}_{j,k}CCTB_IM^{0-17}$ = migration of children according to CCTB data by province/territory of origin j to destination province/territory k;
- ${}_{j}^{CCTB}\mathfrak{R}$ = coverage rate of CCTB program for province/territory of origin j;
- ${}_{j}G$ = adjustment factor for migration propensity of non CCTB/Tax children for province/territory of origin j.

36. See Step 3, Section 7.2.1, for a description of the process to identify interprovincial migrants.

The equation above includes the adjustments for incomplete coverage of CCTB-registered children (Equation 7.2) and for the differences in the propensities of CCTB-registered and non-registered subpopulations to migrate (Equation 7.3).

The coverage rates are calculated on an annual basis as follows:

For each province and territory:

$${}^{\text{CCTB}}_j \mathfrak{R} = \left[\frac{{}^{\text{CCTB}}_j \mathbf{P}^{0-17}}{\text{Dem}_j \mathbf{P}^{0-17}} \right]$$

Equation 7.2:

where

${}^{\text{CCTB}}_j \mathfrak{R}$ = coverage rate of CCTB program for each province/territory j;

${}^{\text{CCTB}}_j \mathbf{P}^{0-17}$ = number of children aged 0 to 17 years registered for the CCTB program in province/territory j;

$\text{Dem}_j \mathbf{P}^{0-17}$ = population estimated by Demography Division of children aged 0 to 17 years in province/territory j.

The differential migration propensity is based on a set of ${}_j \mathbf{G}$ factors. Using tax data, the ${}_j \mathbf{G}$ factors are obtained by dividing the out-migration rates for all children by the out-migration rates of children registered for the CCTB program, as follows:

For each province and territory:

$${}_j \mathbf{G} = \frac{{}_j \text{RateofTax_IM}^{0-17}}{{}_j \text{RateofCCTB_Tax_IM}^{0-17}}$$

Equation 7.3:

$${}_j \mathbf{G} = \frac{\frac{{}_j \text{Tax} \mathbf{I}^{0-17}}{{}_j \text{Tax} \mathbf{P}^{0-17}}}{\frac{{}_j \text{CCTB_Tax} \mathbf{I}^{0-17}}{{}_j \text{CCTB_Tax} \mathbf{P}^{0-17}}}$$

where

${}_j \mathbf{G}$ = adjustment factor for the differential migration propensity by province/territory of origin j;

${}_j \text{RateofTax_IM}^{0-17}$ = out-migration rate of all children from province/territory of origin j, according to income tax data;

${}_j \text{RateofCCTB_Tax_IM}^{0-17}$ = out-migration rate of CCTB-registered children from province/territory of origin j, as modeled from tax data (T1FF);

${}^{\text{Tax}}_j I^{0-17}$	= out-migration of all children from province/territory of origin j, according to income tax data;
${}^{\text{Tax}}_j P^{0-17}$	= SAADD estimate of children in province/territory j based on income tax files;
$\text{CCTB_Tax}_j I^{0-17}$	= out-migration of CCTB-registered children from province/territory j as modeled from tax data (T1FF);
$\text{CCTB_Tax}_j P^{0-17}$	= CCTB-registered children in the population of the income tax files provided by SAADD (T1FF), by province/territory j.

The ${}_j G$ factor will equal one if the migration rate of CCTB-registered children and the migration rate of all children are identical. Otherwise, the value of ${}_j G$ is greater (less) than one if the propensity to migrate of CCTB-registered children is less (greater) than that of all children.

Interprovincial migrant adults

The third factor in adjusting CCTB data involves the calculation of adult migrants. Estimates for adults migrating between provincial/territorial boundaries are obtained through the calculation of F factors. The F factors, calculated from the most recent tax data, are equal to the ratio of the adult migration rate to the child migration rate, by province/territory of origin and destination province/territory, as follows:

For each province and territory of origin and destination:

Equation 7.4:

$${}_{j,k} \bar{F}_{t-3,t} = \left[\frac{{}_{j,k} \text{AvgRateof_IM}^{18+}}{{}_{j,k} \text{AvgRateof_IM}^{0-17}} \right]$$

$${}_{j,k} \bar{F}_{t-3,t} = \frac{\left[\frac{{}^{\text{Tax}}_{j,k} I^{18+}_{t-3,t-2} + {}^{\text{Tax}}_{j,k} I^{18+}_{t-2,t-1} + {}^{\text{Tax}}_{j,k} I^{18+}_{t-1,t}}{{}^{\text{Tax}}_j P^{18+}_{t-2}} \right]}{\left[\frac{{}^{\text{Tax}}_{j,k} I^{0-17}_{t-3,t-2} + {}^{\text{Tax}}_{j,k} I^{0-17}_{t-2,t-1} + {}^{\text{Tax}}_{j,k} I^{0-17}_{t-1,t}}{{}^{\text{Tax}}_j P^{0-17}_{t-2}} \right]}$$

where

- ${}_{j,k} \bar{F}_{t-3,t}$ = average adult estimation factor calculated over three migration periods by province/territory of origin j and destination province/territory k;
- ${}_{j,k} \text{AvgRateof_IM}^{18+}$ = average adult migration rate over three migration periods by province/territory of origin j and destination province/territory k;
- ${}_{j,k} \text{AvgRateof_IM}^{0-17}$ = average child migration rate over three migration periods by province/territory of origin j and destination province/territory k;

${}_{j,k}^{Tax} I^{18+}$	= adult migration by province/territory of origin j and destination province/territory k, according to income tax data for periods (t-3, t-2), (t-2, t-1), (t-1, t);
${}_{j,t-2}^{Tax} P^{18+}$	= adult population estimate for province/territory of origin j at time (t-2) based on income tax data file;
${}_{j,t-2}^{Tax} I^{0-17}$	= child migration by province/territory of origin j and destination province/territory k, according to income tax data for periods (t-3, t-2), (t-2, t-1), (t-1, t);
${}_{j,t-2}^{Tax} P^{0-17}$	= child population estimate for province/territory of origin j at time (t-2) based on income tax data file.

The ${}_{j,k} F$ factors are then applied against child migration rates (based on adjusted CCTB data) to obtain estimates of adult inter-provincial/territorial migration rates. These are multiplied by the provincial/territorial populations to obtain the estimates of the number of adults migrating between provincial/territorial borders, as follows:

For each province and territory:

Equation 7.5:
$${}_{j,k} I^{18+} = \frac{{}_{j,k} I^{0-17}}{\text{Dem}_j P^{0-17}} \times \text{Dem}_j P^{18+} \times {}_{j,k} F_{t-3,t}$$

where

${}_{j,k} I^{18+}$	= estimated number of adults migrating from province/territory of origin j to province/territory k;
$\frac{{}_{j,k} I^{0-17}}{\text{Dem}_j P^{0-17}}$	= migration rate of children (estimated by province/territory of origin j and province/territory of destination k based on the adjusted CCTB data);
$\text{Dem}_j P^{18+}$	= population estimate of adults at province/territory of origin j at the beginning of the period;
${}_{j,k} F_{t-3,t}$	= average adult estimation factor calculated over three migration periods by province/territory of origin j and province/territory of destination k.

Finally, after the estimates of the total number of interprovincial migrant children and adults have been independently calculated by province/territory of origin and destination, the total interprovincial migration estimates are obtained by adding the interprovincial migrants of children and adults, as follows:

For each province and territory:

Equation 7.6:
$${}_{j,k} I = {}_{j,k} I^{0-17} + {}_{j,k} I^{18+}$$

Interprovincial migrant population by age, sex and marital status

For the preliminary series, that is quarterly estimates, we release only the total number of interprovincial migrants by province/territory of origin and destination. However, for the final series, the annual estimates we require a more detailed breakdown of interprovincial and territorial migrants. In order to incorporate interprovincial migration estimates into the population estimation program's annual estimates, the data in each series must therefore be broken down by single year of age, sex and marital status³⁷.

The preliminary quarterly estimate of interprovincial migration shows only the total number of migrants by origin and destination, these are summed up to produce the preliminary annual estimates. A breakdown of these estimates involves an additional stage, during which the totals are broken down by 5-year age groups and sex on the basis of distribution from the most recent final estimates provided by SAADD's T1FF file using data from CRA's tax files. Both the final series of annual interprovincial migration estimates and the preliminary series of derived estimates by age groups and sex, for each province/territory of origin and destination, are then distributed by single year of age and sex using the most recent census information on the one-year mobility question. Lastly, the breakdown is done for each age by marital status, again using the census distributions. This step is carried out for a given year of age and sex for all out-migrants from a province/territory regardless of destination.

37. A description of the methods used to break down the preliminary and final series of interprovincial migrants into single year of age, sex and marital status is given in Bédard and Michalowski, 1997.

Chapter 8

Subprovincial estimates of population

In addition to estimates at the national and provincial/territorial levels, population estimates are produced for subprovincial areas. Annual population estimates are released for census divisions (CDs), census metropolitan areas (CMAs), and economic regions (ERs). Historical series of population estimates at the subprovincial level is available from 1986 onwards. Custom requests for estimates at other levels of geography (e.g., census subdivisions (CSDs) and health regions) are also available.

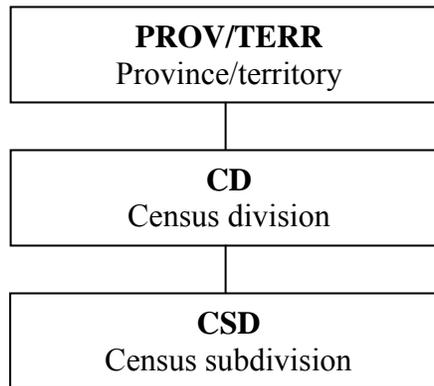
This chapter discusses the estimation methods used for the production of postcensal and intercensal estimates of population for CDs, CMAs and ERs.

8.1 Definition and relevant concepts

A **census division** (CD) is the general term for provincially legislated areas (such as county, *municipalité régionale de comté* and regional district) or their equivalents. Census divisions are intermediate geographic areas between the census subdivision (CSD)³⁸ and the province or territory (Figure 8.1). Census divisions have been established in provincial law to facilitate regional planning, as well as the provision of services that can be more effectively delivered on a scale larger than a municipality. In Newfoundland and Labrador, Manitoba, Saskatchewan, Alberta, Yukon Territory, Northwest Territories and Nunavut, provincial/territorial law does not provide for these administrative geographic areas. Therefore, Statistics Canada, in cooperation with these provinces or territories, has created equivalent areas called census divisions for the purpose of disseminating statistical data. In the Yukon Territory, the census division is equivalent to the entire territory.

38. Census subdivision (CSD) is the general term for municipalities (as determined by provincial legislation) or areas treated as municipal equivalents for statistical purposes (e.g., Indian reserves, Indian settlements and unorganized territories).

Figure 8.1
Standard Geographical Classification (SGC) hierarchy



The general concept of a **census metropolitan area (CMA)** is one of a very large urban area, together with adjacent urban and rural areas which have a high degree of economic and social integration with that urban area. A CMA is formed by one or more adjacent municipalities centered on a large urban area (known as the urban core). The census population count of the urban core is at least 100,000. To be included in the CMA, other adjacent municipalities must have a high degree of integration with the central urban area, as measured by commuting flows derived from census place of work data. A CMA is delineated using adjacent municipalities (CSDs) as building blocks. These CSDs are included in the CMA if they meet at least one of seven delineation rules³⁹. Once an area becomes a CMA, it is retained as a CMA even if the population of its urban core declines below 100,000.

An **economic region (ER)** is a grouping of complete CDs created as a standard geographic unit for analysis of regional economic activity. Within the province of Quebec, economic regions (*régions administratives*) are designated by law. In all other provinces and territories, economic regions are created by agreement between Statistics Canada and the province/territory concerned. Prince Edward Island and the three territories each consist of one economic region. In Ontario, there is one exception where the economic region boundary does not respect census division boundaries: the census division of Halton is split between the ER of Hamilton-Niagara Peninsula and the ER of Toronto.

8.2 Postcensal population estimates of subprovincial areas

8.2.1 Postcensal population estimates for CMAs and CDs

The component method is used to produce population estimates for CMAs and CDs, but figures are adjusted, if necessary, to assure consistency with provincial/territorial totals. Estimates of population are first produced for each province and territory, and then summed to obtain an estimate of the population of Canada. The estimates at the subprovincial level are adjusted proportionately to coincide with annual estimates of the total population for provinces and territories.

39. Refer to **2001 Census Dictionary**, catalogue no. 92-378-X, pages 229-235 for a description of the seven delineation rules for CMAs.

The component method formula used in estimating total CMA and CD populations is expressed as follows:

For every subprovincial area:

Equation 8.1:

$$P_{(t+1)} = P_t + B_{(t,t+1)} - D_{(t,t+1)} + I_{(t,t+1)} - (E_{(t,t+1)} + \Delta TE_{(t,t+1)}) + RE_{(t,t+1)} + \Delta NPR_{(t,t+1)} + \Delta N_{(t,t+1)}^{inter} + \Delta N_{(t,t+1)}^{intra}$$

where, for each subprovincial area:

$(t, t+1)$ = interval between the date of the previous reference period, time t , and the reference date of the estimate, time $t+1$;

$P_{(t+1)}$ = estimate of the population at time $t+1$;

P_t = base population at time t (census adjusted for net census undercoverage or most recent estimate);

$B_{(t,t+1)}$ = number of births between time t and $t+1$;

$D_{(t,t+1)}$ = number of deaths between time t and $t+1$;

$I_{(t,t+1)}$ = number of immigrants between time t and $t+1$;

$E_{(t,t+1)}$ = number of emigrants between time t and $t+1$;

$\Delta TE_{(t,t+1)}$ = net temporary emigration between time t and $t+1$;

$RE_{(t,t+1)}$ = number of returning emigrants between time t and $t+1$;

$\Delta NPR_{(t,t+1)}$ = net non-permanent residents between time t and $t+1$ (can be either positive or negative);

$\Delta N_{(t,t+1)}^{inter}$ = net interprovincial migration between time t and $t+1$ (can be either positive or negative);

$\Delta N_{(t,t+1)}^{intra}$ = net intraprovincial migration between time t and $t+1$ (can be either positive or negative).

In the estimation of migration, an additional component, subprovincial migration or intraprovincial migration, (i.e., migration within provinces/territories and across subprovincial areas) is necessary.

8.2.2 Postcensal population estimates for ERs

A different method is used to produce population estimates for ERs: the CD aggregation method. First, the ERs are defined in terms of CDs using Standard Geographical Classification (SGC) specifications. When the geographic delineation of the CDs matches that of the ER, no adjustment is required; the population estimates for the CDs that make up the ER are simply added together.

However, when the geographic delineation of the CDs does not match that of the ER, or when a CD is in more than one ER, allocation of the CD's demographic components is prorated on the basis of its proportion of each ER's population. The proportions are referred to as conversion factors. They are calculated using the most recent census counts.

Subsequently, all demographic components, with the exception of intraprovincial migration, initially measured at the CD level can be allocated to each ER. In other words, the population and demographic components of ERs can be estimated by aggregating the CD data based on the ERs' geographic delineation.

The CD aggregation method is not an appropriate method to estimate the components of intraprovincial migration for ERs as it overestimates the number of in-migrants and out-migrants. In-migrants to a given CD from another CD in the same ER should not be counted as the migration occurred within the ER's boundaries. Thus, they are false in-migrants. The same is true for out-migrants from one CD to another CD in the same ER. Again, they are false out-migrants. However, combining the in-migration and out-migration figures as derived by the CD aggregation method produces a consistent result since the false in-migrants and false out-migrants cancel out. Thus, only the net intraprovincial migration of ERs can be estimated accurately by the CD aggregation method, and due to the reasons stated above, the numbers of intraprovincial in-migrants and out-migrants are not available for ERs.

Special treatment for preliminary postcensal estimates for Quebec and British Columbia

A different method is used to calculate preliminary postcensal population estimates for CDs and CMAs in Quebec and British Columbia. For Quebec, the total population estimates produced by the *Institut de la statistique du Québec* (ISQ) are used. Those estimates are based on data from the insured persons registration file of the *Régie de l'assurance-maladie du Québec*. They are benchmarked to correspond to Demography Division's estimate of Quebec's total population. The same approach is followed for the ISQ's ER-level estimates, which are derived from its CD-level estimates.

For British Columbia, preliminary postcensal estimates are computed using CD and CMA growth rates provided by British Columbia's statistics agency, BC STATS, for total population only. To produce new population estimates for each CD and CMA, the rates are applied to Demography Division's estimates of total population for the previous year. Prorating is used to ensure that the province's total population released by Demography Division equals the sum of the populations of the subprovincial areas. The British Columbia population estimates used to calculate the growth rates are produced using the Difference-Correlation method, a regression model based on residential electrical (Hydro) connections and Old Age Security data as symptomatic indicators.

8.2.3 Subprovincial postcensal population estimates by age and sex

Postcensal population estimates by age and sex for CDs, CMAs and ERs are produced by applying the component method to each age-sex cohort in the base population, whereby the population is aged from year to year and the components are tabulated according to age and sex cohorts. Chapter 9 describes the application of the cohort component approach in details.

Special treatment for preliminary postcensal estimates by age and sex for Quebec and British Columbia

A different method is used to calculate preliminary postcensal population estimates by age and sex for CDs, CMAs and ERs in Quebec. The population estimates by age and sex produced by the *Institut*

de la statistique du Québec (ISQ) are used. These estimates are based on data from the insured persons registration file of the *Régie de l'assurance-maladie du Québec*. They are benchmarked to correspond to Demography Division's total estimates by age and sex for Quebec.

For British Columbia, the estimates by age and sex are produced by applying the age-sex distribution derived by the component method to the total population estimates for CMAs and CDs obtained by applying the growth rates, supplied by BC STATS. Two-way raking is used to ensure that the estimates are consistent with the provincial totals by age and sex. The same approach is followed for the ER-level estimates, which are derived from the CD-level estimates.

8.2.4 Levels of estimates

For Quebec and British Columbia, methods described in the previous sections are used only for preliminary postcensal estimates. For updated and final postcensal estimates, the component method is used.

For the other provinces and territories, the difference between preliminary and final postcensal population estimates lies in the timeliness of the components. When all the components are preliminary, the population estimate is described as preliminary postcensal (PP). When they are all final, the estimate is referred to as final postcensal (PD). Any other combination of levels is referred as updated postcensal (PR).

8.2.5 Base population and components of population growth

Base population

A full description of the methodology for deriving the postcensal base population is described in Chapter 2. In the population estimation program, the base populations for subprovincial areas are derived from the quinquennial censuses between 1986 and 2001, whereas the base populations for provinces and territories are derived from the 1971 to 2001 censuses. Population counts at both the provincial/territorial and subprovincial levels are subjected to the same adjustment procedures as outlined in Chapter 2, unless otherwise noted. Coverage studies do not provide estimates of net undercoverage for CDs and CMAs. In order to obtain estimates at subprovincial levels, provincial rates of net undercoverage by single years of age and sex are applied to all geographic regions in the province.

Prior to generating the population estimates for each component, the total province/territorial population counts and the counts for CDs and CMAs are adjusted, if necessary, to assure consistency between the two sets of figures. To adjust these data, two adjustment procedures are used: prorating and two-way raking. Prorating is used to ensure that the province's total population equals the sum of the populations of the subprovincial areas. The technique adjusts the total population counts to the components of subprovincial areas. It distributes the difference between the province's total population and the sum of the population of the subprovincial areas proportionally across those areas.

Two-way raking is an iterative procedure whereby differences are simultaneously adjusted in assuming a consistency between:

1. the sum of subprovincial areas and provincial/territorial population totals, and
2. the sum across subprovincial areas for specific age/sex categories and the provincial/territorial distribution by age and sex.

Since CD and CMA boundaries do not remain stable over time, component data are adjusted to respect the boundaries defined in the 2001 Census. This ensures a stable base population for all estimates.

Births and deaths

In addition to the national and provincial/territorial data on births and deaths, the numbers are also available for CSDs and CDs from the Vital Statistics databases of Statistics Canada's Health Statistics Division. Births and deaths data for CMAs are derived by Demography Division using data at lower levels of geography. Where appropriate, the estimates of births and deaths are categorized as final. To ensure their consistency, the estimates are subsequently benchmarked to the provincial/territorial totals using two-way raking.

When no data are available for births and deaths, subprovincial estimates are produced by disaggregating the preliminary provincial/territorial estimates on the basis of the most recent subprovincial distribution derived from Health Statistics Division's Vital Statistics. In such case, the estimates of births and deaths are categorized as preliminary. To ensure their consistency, the estimates are subsequently benchmarked to the provincial/territorial totals using two-way raking.

Immigration

The distribution of the number of immigrants by CD and CMA cannot be obtained from the CIC files as with the provincial/territorial level estimates, because the geographic coding is specific to CIC and is not directly convertible to the coding used by Statistics Canada. Subprovincial estimates are produced by disaggregating the preliminary and final provincial/territorial estimates based on the most recent subprovincial distribution derived from SAADD's T1 Family File (T1FF)⁴⁰. The data are available only by broad age groups and must be disaggregated by sex and single years of age based on the provincial distribution. To ensure their consistency, they are subsequently benchmarked to the provincial/territorial totals using two-way raking.

Non-permanent residents (NPRs)

At the subprovincial level, there are no reliable administrative data available to estimate NPRs. To compensate for the lack of data, the provincial/territorial estimates by age and sex are disaggregated by subprovincial area on the basis of the subprovincial distribution in the most recent census. To ensure their consistency, they are subsequently benchmarked to the provincial/territorial totals using two-way raking.

40. The T1 Family File is derived from the Canada Revenue Agency's T1 file by Small Area and Administrative Data Division (SAADD) of Statistics Canada.

Emigration

As in the case of immigrants, the number of emigrants at the subprovincial level is derived from SAADD's T1FF. The estimates are available only by broad age groups and must be disaggregated by sex and single years of age based on the provincial/territorial distribution. To ensure their consistency, they are subsequently benchmarked to the provincial/territorial totals using two-way raking.

Net temporary emigration

At the subprovincial level, provincial/territorial net temporary emigration is disaggregated on the basis of the age-sex distribution of subprovincial emigrants. To ensure their consistency, the estimates are subsequently benchmarked to the provincial/territorial totals using two-way raking.

Returning emigrants

The provincial/territorial numbers of returning emigrants are disaggregated by subprovincial level based on the age-sex distribution of subprovincial immigrants. To ensure their consistency, they are subsequently benchmarked to the provincial/territorial totals using two-way raking.

Interprovincial migration

The Canada Child Tax Benefit administrative files are used to produce preliminary estimates of interprovincial migration, while data from SAADD's T1FF are used to produce the final estimates for subprovincial areas. Interprovincial migration for CDs and CMAs by broad age groups and sex are generated by disaggregating the preliminary and final provincial/territorial estimates on the basis of SAADD's T1FF. The data are disaggregated into single years of age using distributions from the 2001 Census one-year mobility question. To ensure their consistency, they are subsequently benchmarked to the provincial/territorial totals using two-way raking.

Intraprovincial migration

The estimates of intraprovincial migration by broad age group and sex are derived from SAADD's T1FF for both preliminary and final estimates. The estimates by broad age group and sex are disaggregated into single years of age using distributions from the 2001 Census one-year mobility question.

8.3 Intercensal population estimates of subprovincial areas

Intercensal estimates for the 1986 to 2001 period at the census division (CD), census metropolitan area (CMA), and economic region (ER) levels are produced much in the same manner as intercensal estimates at the provincial/territorial level (for information on the methods, refer to Chapter 1). However, the production of these estimates for CDs, CMAs and ERs additionally require that changes in census geography over time are accounted for. The base population and components of population growth from 1986 to 2001 had to be adjusted according to the 2001 Census boundaries. This was done

by applying conversion factors based on the population of the 2001 Census subdivisions. For most provinces, there were only minor corrections to CDs, CMAs and ERs.

As for the provinces/territories, the error of closure is calculated for each subprovincial area by comparing the postcensal estimates with the most recent censal estimates, by age and sex. The resultant error of closure is then distributed linearly to the postcensal population estimates for those years between the previous and most recent censuses. When summed, the series of intercensal estimates by age and sex for subprovincial areas differ slightly from the intercensal estimates of total population aggregated to the provincial and territorial level. In order to ensure consistency, the subprovincial estimates by age and sex are adjusted to the provincial estimates using two-way raking.

Chapter 9

Population estimates by age, sex and marital status

Analysis of population by age and sex is a fundamental aspect of most demographic studies. The age and sex structure of the population varies with time and place, while at the same time demographic behaviour is often a function of age and sex. For example, mortality rates are much higher in the older age groups. High migration rates are associated with young adults, as they move for personal and economic reasons. Population estimates by age and sex are widely used by other divisions of Statistics Canada. For example, these estimates are used in the calculation of employment and unemployment rates and crime rates, which tend to vary according to age and sex distributions. The addition of marital status builds the foundation for studying other demographic phenomena such as marriage and divorce rates and changes in family structure. Government and private sector planning and policies are largely driven by the age, sex and marital status profiles of certain populations.

Population estimates by age and sex are available at national, provincial/territorial, census division, census metropolitan area and economic region levels. The more detailed breakdown of these estimates by marital status is available only at the national and provincial/territorial levels. This chapter presents the methods used to produce population estimates disaggregated by age, sex and marital status.

9.1 Postcensal population estimates by age and sex, Canada, provinces and territories

9.1.1 Data sources and relevant concepts

Postcensal estimates of population by age and sex are produced using the cohort component approach. This is similar to the component method as used in the production of total population estimates, although additional data are required in its application. The data required for the cohort component method are related to demographic events (deaths, immigration, non-permanent residents, internal migration, emigration, returning emigrants, temporary emigration) that can be directly linked to persons belonging to the same birth cohort (i.e., persons having been born during the same period or year). Different components require unique treatment, according to the nature of the data used to generate the estimates, and their respective chapters elaborate upon the manner in which the estimate for each component is distributed by age and sex.

The data sources used in the production of the population estimates by age and sex are as follows:

- Births and deaths using Vital Statistics;
- Immigration and non-permanent residents using data from Citizenship and Immigration Canada (CIC);
- Emigration are distributed by age and sex using data from the Office of Immigration Statistics, U.S. Department of Homeland Security, Small Area and Administrative Data Division (SAADD) and CIC;
- Net temporary emigration are distributed by age and sex using emigration distribution;

- Returning emigrants are distributed by age and sex using the most recent census data on Canadians who were living abroad one year prior to the census ;
- Interprovincial migrants are distributed by age and sex based on data from SAADD's broad age/sex groupings and then broken down by single age within those categories based on the most recent census information on one-year mobility question.

9.1.2 Levels of estimates

The difference between preliminary and final postcensal estimates lies in the timeliness of the components. When all the components are preliminary, the estimate is described as preliminary postcensal (PP). When they are all final, the estimate is referred to as final postcensal (PD). Any other combination of levels is referred to as updated postcensal (PR).

9.1.3 Methods of estimation

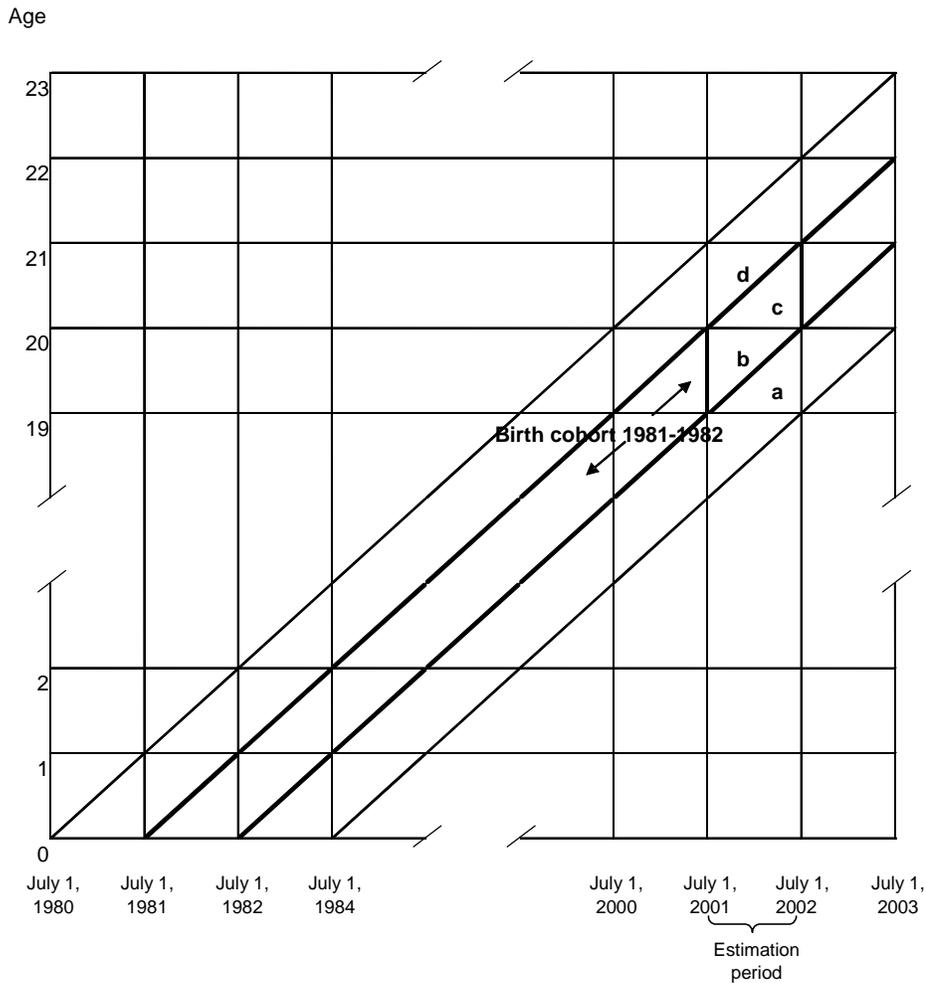
Postcensal estimates of population by age and sex are produced using the cohort component approach. This approach requires a slight modification of the component approach described in Chapter 1, but the overall principles are the same.

Annual estimates

Estimates of population by age and sex are published annually with July 1st as the reference date. To calculate these estimates, birth cohorts (those persons born during the same year) for both males and females separately, are used. The cohort component approach factors in the ageing of the cohorts over time. For example, persons aged 19 one year will be 20 years old the following year. The data required for the cohort component method include demographic events such as deaths, immigration, emigration, net temporary emigration, returning emigration, interprovincial migration and non-permanent residents that can be directly linked to persons belonging to the same birth and sex cohorts.

Demographers use a tool called a Lexis diagram (Figure 9.1) to aid in the linking of events to specific cohorts. Time is located on the horizontal axis (abscissa), while the vertical axis (ordinate) represents age. Specific cohorts are identified by the diagonals (lifelines) that cross the **diagram**. Using the cohort component approach, demographic events are organized to follow these lifelines.

Figure 9.1
Transition from a distribution of demographic events by age and period to a distribution by age and birth cohort



Take, for example, those aged 19 as of July 1, 2001, who belong to the cohort born between July 1, 1981 and June 30, 1982 (inclusive). The demographic events experienced by this cohort during the estimation period are represented by triangles “b” and “c”.

The equations for estimating annual population by single years of age and sex, by the cohort component method (at the national and provincial/territorial levels) are as follows:

For each sex, by province/territory:

Age 0:

Equation 9.1:

$$P_{(t+1)}^0 = B_{(t,t+1)} - D_{(t,t+1)}^0 + I_{(t,t+1)}^0 - (E_{(t,t+1)}^0 + \Delta TE_{(t,t+1)}^0) + RE_{(t,t+1)}^0 + \Delta NPR_{(t,t+1)}^0 + \Delta N_{(t,t+1)}^0$$

Ages 1 through 89:

Equation 9.2:

$$P_{(t+1)}^{(a+1)} = P_t^a - D_{(t,t+1)}^a + I_{(t,t+1)}^a - (E_{(t,t+1)}^a + \Delta TE_{(t,t+1)}^a) + RE_{(t,t+1)}^a + \Delta NPR_{(t,t+1)}^a + \Delta N_{(t,t+1)}^a$$

Ages 90+:

Equation 9.3:

$$P_{(t+1)}^{90+} = P_t^{89+} - D_{(t,t+1)}^{89+} + I_{(t,t+1)}^{89+} - (E_{(t,t+1)}^{89+} + \Delta TE_{(t,t+1)}^{89+}) + RE_{(t,t+1)}^{89+} + \Delta NPR_{(t,t+1)}^{89+} + \Delta N_{(t,t+1)}^{89+}$$

where

- (t, t + 1) = interval between the date of the previous reference period, time t, and the reference date of the estimate, time t+1;
- a = age at time t;
- $P_{(t+1)}^0$ = postcensal estimate for persons of age less than one year at time (t+1);
- $P_{(t+1)}^{(a+1)}$ = postcensal estimate for persons of age (a+1) at time (t+1), for persons aged 1 to 89 years;
- $P_{(t+1)}^{90+}$ = postcensal estimate for persons aged 90 years and over at time (t+1);
- P_t^a = base population at time t (census adjusted for net census undercoverage, or most recent estimates) at age a;
- $B_{(t,t+1)}$ = number of births between time t and t+1;
- $D_{(t,t+1)}$ = number of deaths between time t and t+1;
- $I_{(t,t+1)}$ = number of immigrants between time t and t+1;
- $E_{(t,t+1)}$ = number of emigrants between time t and t+1;
- $\Delta TE_{(t,t+1)}$ = net temporary emigration between time t and t+1;
- $RE_{(t,t+1)}$ = number of returning emigrants between time t and t+1;
- $\Delta NPR_{(t,t+1)}$ = net non-permanent residents between time t and t+1;
- $\Delta N_{(t,t+1)}$ = net interprovincial migration between time t and t+1.

9.2 Intercensal population estimates by age and sex, Canada provinces and territories

Intercensal population estimates for reference dates between two censuses are produced following each census. They reconcile previous postcensal estimates with the new census counts. Like the total population by province/territory, intercensal population by age and sex are adjusted by distributing the error of closure uniformly across the age-sex cohorts.

9.3 Subprovincial postcensal and intercensal estimates by age and sex

Postcensal population estimates by age and sex for census divisions (CDs), census metropolitan areas (CMAs) and economic regions (ERs) are produced by applying the component method to each age-sex cohort in the base population, whereby the population is aged from year to year and the components are tabulated according to age and sex cohorts.

Special methods for preliminary postcensal estimates by age and sex are applied for Quebec and British Columbia. These methods and the approach used to derive intercensal estimates by age and sex at subprovincial levels are described in Chapter 8.

9.4 Population estimates by age, sex and marital status, Canada, provinces and territories

There are two series of population estimates by marital status, the main difference between them being the treatment of persons living in “common law” unions. One of them is the series of estimates by legal marital status, i.e., a person’s conjugal status under the law (for example, single, married, widowed or divorced). On the basis of this definition, people living “common law” are categorized by their legal marital status. If a person has never “married” and is living “common law”, he or she is regarded as “single” under this definition.

The other is the series of estimates by marital status, i.e., a person’s de facto conjugal status. For example, a person who reports being legally “widowed” and is living with another person as a couple but is not married to that person will be counted as “common law” in the marital status series and “widowed” in the legal marital status series.

Separate estimates for legal and de facto marital statuses at the national, provincial and territorial levels are available from 1991 onwards. Estimates of marital statuses are not produced for subprovincial levels. Postcensal estimates of the population by age, sex and marital status dating back to 1971 essentially reflected a combination of legal and de facto marital statuses.

9.4.1 Data sources and relevant concepts

Marital status refers to the conjugal status of a person. In demographic estimates, a distinction is made between “legal” marital status and marital status. The distinction between the two definitions lies in the concept of who is considered married. In the discussion of legal marital status, a person’s marital status is determined by law. Common-law partners are not legally married to each other, thus are considered single, divorced or widowed according to their legal marital status. Separated couples are considered married under both concepts.

The following definitions represent those used by Statistics Canada for legal marital status and marital status, respectively.

Legal marital status refers to the marital status of the person under the law. Estimates are presented in the following categories: single, legally married, widowed or divorced.

Single: Persons who have never been married or persons whose marriage has been annulled and have not remarried. All persons aged less than 15 are considered as never married (single);

Legally married: Persons whose spouse is living, unless a divorce has been obtained. Persons separated are also included in this category;

Widowed: Persons who have lost their spouse through death and who have not remarried;

Divorced: Persons who have obtained a divorce and who have not remarried.

Marital status indicates the conjugal arrangement of a person. Estimates are presented in the following categories: single, married (including persons living common-law and persons who are separated), widowed or divorced. Common-law status refers to whether the person aged 15 or over is living with a person of the opposite sex or of the same sex as a couple but is not legally married to that person. It includes situations where the members of such a couple are living apart temporarily because of illness, work or school.

Single: Persons who have never been married, or persons whose marriage has been annulled and who have not remarried, and do not live common-law. All persons aged less than 15 are considered as never married (single);

Married (including persons living common-law and persons who are separated): Persons whose spouse is living, unless a divorce has been obtained;

Widowed: Persons who have lost their spouse through death and who have not remarried, and who do not live common-law;

Divorced: Persons who have obtained a divorce and who have not remarried, and who do not live common-law.

To produce population estimates by age, sex and marital status, data about events resulting in a change of marital status (for example, marriage, divorce, death of a spouse) are also required. These data are available from the Vital Statistics database maintained by Health Statistics Division of Statistics Canada. The database was created in collaboration with provincial and territorial ministries as well as the federal department of Justice Canada. This collaboration allows Statistics Canada to obtain data on marriage certificates registered by the provinces/territories for legal purposes as well as data related to divorces from the Central Divorce Registry⁴¹ maintained by Justice Canada.

9.5 Postcensal population estimates by legal marital status

9.5.1 Methods of estimation

Population estimates by legal marital status are produced by the cohort component method. This method takes into account events that result in a change of legal marital status. All persons under the age of 15 years are assumed to be single. For persons aged 15 years and over, estimates by single year of age are prepared for four subpopulations: single, legally married, divorced and widowed, as previously defined for legal marital status. These estimates are based on the census counts, adjusted to July 1st, as well as for net census undercoverage. The non-permanent residents were added to the census populations prior to 1991 since they became part of the census universe only from 1991 on.

The component method formulas used to estimate the four legal marital status subpopulations are as follows:

For each sex and age (15 years and older), by province/territory:

41. The Central Divorce Registry is a database containing all the information related to divorce decrees granted in Canadian courts.

Single:*Equation 9.4:*

$$P_{(t+1)}^{nm(\text{legal}),(a+1)} = P_t^{nm,a} - D_{(t,t+1)}^{nm,a} + I_{(t,t+1)}^{nm,a} - (E_{(t,t+1)}^{nm,a} + \Delta TE_{(t,t+1)}^{nm,a}) + RE_{(t,t+1)}^{nm,a} + \Delta NPR_{(t,t+1)}^{nm,a} + \Delta N_{(t,t+1)}^{nm,a} - \text{Mar}_{(t,t+1)}^{nm,a}$$

Legally married:*Equation 9.5:*

$$P_{(t+1)}^{mar(\text{legal}),(a+1)} = P_t^{mar,a} - D_{(t,t+1)}^{mar,a} + I_{(t,t+1)}^{mar,a} - (E_{(t,t+1)}^{mar,a} + \Delta TE_{(t,t+1)}^{mar,a}) + RE_{(t,t+1)}^{mar,a} + \Delta NPR_{(t,t+1)}^{mar,a} + \Delta N_{(t,t+1)}^{mar,a} + \text{Mar}_{(t,t+1)}^a - \text{Veu}_{(t,t+1)}^a - \text{Div}_{(t,t+1)}^a$$

Divorced:*Equation 9.6:*

$$P_{(t+1)}^{v(\text{legal}),(a+1)} = P_t^{v,a} - D_{(t,t+1)}^{v,a} + I_{(t,t+1)}^{v,a} - (E_{(t,t+1)}^{v,a} + \Delta TE_{(t,t+1)}^{v,a}) + RE_{(t,t+1)}^{v,a} + \Delta NPR_{(t,t+1)}^{v,a} + \Delta N_{(t,t+1)}^{v,a} - \text{Mar}_{(t,t+1)}^{v,a} + \text{Div}_{(t,t+1)}^a$$

Widowed:*Equation 9.7:*

$$P_{(t+1)}^{w(\text{legal}),(a+1)} = P_t^{w,a} - D_{(t,t+1)}^{w,a} + I_{(t,t+1)}^{w,a} - (E_{(t,t+1)}^{w,a} + \Delta TE_{(t,t+1)}^{w,a}) + RE_{(t,t+1)}^{w,a} + \Delta NPR_{(t,t+1)}^{w,a} + \Delta N_{(t,t+1)}^{w,a} - \text{Mar}_{(t,t+1)}^{w,a} + \text{Veu}_{(t,t+1)}^a$$

where for each province and territory:

a	= age at time t;
(t,t+1)	= interval between the date of the previous reference period, time t, and the reference date of the estimate, time t+1;
legal	= legal marital status;
nm	= never married;
mar	= legally married;
v	= divorced;
w	= widowed;
P_{t+1}^{a+1}	= estimate of the population at age a+1 at time t+1;
P_t^a	= base population at time t (census adjusted for NCU, or most recent estimates) at age a;
$B_{(t,t+1)}$	= number of births between time t and t+1;
$D_{(t,t+1)}$	= number of deaths between time t and t+1;
$I_{(t,t+1)}$	= number of immigrants between time t and t+1;
$E_{(t,t+1)}$	= number of emigrants between time t and t+1;
$\Delta TE_{(t,t+1)}$	= net temporary emigration between time t and t+1;
$RE_{(t,t+1)}$	= number of returning emigrants between time t and t+1;
$\Delta NPR_{(t,t+1)}$	= net non-permanent residents between time t and t+1;
$\Delta N_{(t,t+1)}$	= net interprovincial migration between time t and t+1;
Mar _(t,t+1)	= number of marriages between time t and t+1;

$Div_{(t,t+1)}$ = number of divorces between time t and t+1;

$Ve_{(t,t+1)}$ = new widowhoods between time t and t+1.

The components of population growth are disaggregated by legal marital status as follows:

1. **deaths** by age and sex are disaggregated by legal marital status using the latest death statistics published by Health Statistics Division;
2. **interprovincial migration** is disaggregated by legal marital status using the counts from the census used in calculating the base population (one-year mobility question);
3. **immigration and non-permanent residents** are disaggregated by legal marital status using the estimates derived from the files of Citizenship and Immigration Canada's Field Operational Support System (FOSS);
4. **emigration and net temporary emigration** is disaggregated by legal marital status using the latest available data from the Office of Immigration Statistics, U.S. Department of Homeland Security;
5. **number of returning emigrants** is disaggregated by legal marital status using the population counts from the census used in the base population (one-year mobility question);
6. **marriages** are disaggregated by age, sex and legal marital status using the latest marriage statistics published by Health Statistics Division. The totals are estimated beforehand by Demography Division;
7. **divorces** are disaggregated by age, sex and legal marital status using the latest Justice Canada divorce statistics published by Health Statistics Division. The totals are estimated beforehand by Demography Division;
8. the difficulty with deaths of married persons is that the age of the surviving spouse is not recorded in the Vital Statistics Registry. Consequently, the age of the new widow or widower must be determined indirectly. **New widowhoods** are derived from estimates of deaths of married persons and husband-wife families based on the population counts from the census used in calculating the base population.

New widowhoods

Using the age distribution of deaths to married males and females, compiled by age⁴² group, the surviving spouse is assigned to a five-year age group. Assignment is based on the distribution of legally married husband-wife families by age group of husband and age group of wife from the last census (see Table 9.1).

42. Age at the beginning of the estimation period.

Table 9.1**Percent distribution of husband-wife families by age group of wife and age group of husband according to the legal marital status concept: Ontario, 2001 Census**

Age group of husband	Age group of wife (years)								
	15 to 19 (1)	20 to 24 (2)	25 to 29 (3)	30 to 34 (4)	35 to 39 (5)	40 to 44 (6)	45 to 49 (7)	50 to 54 (8)	55 to 59 (9)
15 to 19	19.0	13.6	12.2	5.7	14.2	8.9	12.8	2.0	0.8
20 to 24	3.4	65.4	23.8	3.6	0.9	0.6	0.9	0.7	0.1
25 to 29	0.3	17.7	64.1	15.1	2.0	0.4	0.2	0.1	0.1
30 to 34	0.1	3.3	28.6	53.2	12.1	2.1	0.4	0.1	0
35 to 39	0	0.7	6.1	31.6	49.5	10.0	1.6	0.4	0.1
40 to 44	0	0.2	1.5	7.8	35.5	45.4	8.0	1.3	0.2
45 to 49	0	0	0.3	1.9	9.5	37.1	43.0	7.1	0.9
50 to 54	0	0	0.1	0.6	2.6	10.9	37.3	41.8	5.6
55 to 59	0	0	0.1	0.3	0.9	3.1	11.5	41.7	36.3
60 to 64	0	0	0	0.2	0.4	1.1	3.3	13.7	40.5
65 to 69	0	0	0	0.1	0.2	0.5	1.2	4.3	12.8
70 to 74	0	0	0	0.1	0.1	0.2	0.5	1.5	4.3
75 to 79	0	0	0	0	0.1	0.1	0.2	0.6	1.4
80 to 84	0	0.1	0	0	0	0.1	0.1	0.4	0.5
85 to 89	0	0	0	0.1	0.1	0.2	0.1	0.2	0.3
90+	0.2	0	0.7	0.5	1.0	0.5	0	0	0.4
	60 to 64 (10)	65 to 69 (11)	70 to 74 (12)	75 to 79 (13)	80 to 84 (14)	85 to 89 (15)	90+ (16)	Total (17)	
15 to 19	1.4	3.9	1.4	4.1	0	0	0	100.0	
20 to 24	0.3	0	0.1	0.1	0.1	0	0	100.0	
25 to 29	0	0	0	0	0	0	0	100.0	
30 to 34	0.1	0	0	0	0	0	0	100.0	
35 to 39	0	0	0	0	0	0	0	100.0	
40 to 44	0.1	0	0	0	0	0	0	100.0	
45 to 49	0.2	0	0	0	0	0	0	100.0	
50 to 54	0.9	0.2	0	0	0	0	0	100.0	
55 to 59	5.2	0.7	0.2	0	0	0	0	100.0	
60 to 64	34.4	5.3	0.8	0.2	0	0	0	100.0	
65 to 69	38.9	34.9	5.9	1.1	0.1	0	0	100.0	
70 to 74	13.4	36.5	35.7	6.7	0.9	0.1	0	100.0	
75 to 79	3.9	13.3	39.9	35.0	5.0	0.4	0.1	100.0	
80 to 84	1.8	4.9	17.8	42.3	28.1	3.6	0.3	100.0	
85 to 89	1.0	2.2	8.5	25.3	38.7	21.1	2.2	100.0	
90+	0.5	0.6	5.2	12.2	24.3	36.3	17.6	100.0	

Note: The percentages in this table are rounded to sum exactly to 100.0.

Source: 2001 Census data

The number of new widow(er)s by age group ($a, a + 4$), for $a = 15, 20, \dots, 85$, is estimated as follows for each province/territory:

$$\text{Equation 9.8: } \text{Ve}u^{(a,a+4)} = \sum_{(b=15)}^{90+} D^{mar,(b,b+4)} \times \Phi^{(a,a+4)}$$

with

$$\text{Equation 9.9: } \Phi^{(a,a+4)} = \frac{F^{(a,a+4)}}{\sum_{(a=15)}^{90+} F^{(a,a+4)}}$$

where

- $V_{eu}^{(a,a+4)}$ = estimated number of new widow(er)s in age group $(a, a + 4)$;
 $D^{mar,(b,b+4)}$ = number of deaths of married persons (male or female) in age group $(b, b + 4)$;
 $\Phi^{(a,a+4)}$ = proportion of husband-wife families by five-year age group of husbands (wives) cross-classified by five-year age group of wives (husbands) $(a, a + 4)$;
 $F^{(a,a+4)}$ = number of husband-wife families by five-year age group of husbands (wives) cross-classified by five-year age group of wives (husbands) $(a, a + 4)$.

Table 9.2 provides an illustration of the above formulae. A sample calculation of the incidence of widowhood among females in the 45-49 year age group for the province of Ontario in 2001/2002 is presented. The percent distribution of husbands with a 45 to 49 year old wife is taken from column (7) in Table 9.1. This is applied to the 2001/2002 distribution of male deaths (column (1) in Table 9.2), and the results summed to estimate the number of new widows aged 45 to 49 years, as presented in column (3). To determine the number of widows in another age group, the appropriate distribution from Table 9.1 is substituted for column (2) of Table 9.2, and the same procedure is followed.

Table 9.2
Calculation of the incidence of widowhood according to the legal marital status concept: An example for females aged 45 to 49 years, Ontario, 2001/2002

Age group of husband	Deaths of married males	Percent distribution of husbands with wives aged 45 to 49 years	Estimates of new widows aged 45 to 49 years
	(1)	(2)	(3) = [(1) X (2)] / 100
15 to 19	0	12.8	0.000
20 to 24	10	0.9	0.086
25 to 29	54	0.2	0.091
30 to 34	143	0.4	0.601
35 to 39	226	1.6	3.699
40 to 44	430	8.0	34.189
45 to 49	695	42.9	298.397
50 to 54	1,146	37.3	427.597
55 to 59	1,474	11.5	169.526
60 to 64	2,099	3.3	68.605
65 to 69	2,907	1.2	35.143
70 to 74	4,039	0.5	18.732
75 to 79	4,613	0.2	10.373
80 to 84	3,864	0.1	3.982
85 to 89	2,402	0.1	1.766
90+	960	0.0	0.258
Total			1,073.045

Rounded to:
1,073

Source: Statistics Canada, Demography Division

These results by five-year age group are then disaggregated into single years of age using Sprague's multipliers⁴³. This distribution of surviving spouses represents the spouse's age at the beginning of the reference period.

9.5.2 Levels of estimates

The difference between preliminary and final postcensal estimates of the population by legal marital status lies in the timeliness of the components. When all the components are preliminary, the estimate is described as preliminary postcensal (PP). When they are all final, the estimate is referred to as final postcensal (PD). Any other combination of levels is referred to as updated postcensal (PR).

9.6 Postcensal population estimates by marital status

9.6.1 Methods of estimation

Since there are no reliable sources of data on annual formations and dissolutions of common-law unions, an alternative method is used to produce postcensal estimates by marital status. The number of persons living common-law is estimated using estimates by legal marital status and proportions derived from the census adjusted for net census undercoverage. The proportion of never-married persons living common-law is obtained by dividing the number of never-married persons living common-law by the total number of never-married persons. The same procedure is used to calculate the proportion of divorced and widowed persons living common-law. The marital status proportions are generated for each age, sex, province and territory.

The derived proportions are then multiplied by the number of never-married, divorced and widowed persons respectively to produce the number of persons living common-law. The number of persons living common-law by marital status is then subtracted from the estimates by legal marital status and added to the married category to produce the population by marital status.

As for the estimates by age and sex, the base populations by marital status are adjusted to July 1st and for net census undercoverage. The non-permanent residents are added to the census populations prior to 1991 since they became part of the census universe only from 1991 on.

Postcensal population estimates by marital status for a given age and sex can be expressed as follows:

For each province/territory, by sex:

Single:

Equation 9.10:

$$P_{(t+1)}^{nm(ms),a} = P_{(t+1)}^{nm(legal),a} - P_{(t+1)}^{nm(cl),a}$$

43. Sprague's multipliers are interpolation coefficients used to subdivide data. For a detailed description of the use of Sprague's multipliers, see Shryock et al. (1976).

Married:

Equation 9.11:

$$P_{(t+1)}^{\text{mar}(ms),a} = P_{(t+1)}^{\text{mar}(legal),a} + P_{(t+1)}^{\text{nm}(cl),a} + P_{(t+1)}^{\text{w}(cl),a} + P_{(t+1)}^{\text{v}(cl),a}$$

Divorced:

Equation 9.12:

$$P_{(t+1)}^{\text{v}(ms),a} = P_{(t+1)}^{\text{v}(legal),a} - P_{(t+1)}^{\text{v}(cl),a}$$

Widowed:

Equation 9.13:

$$P_{(t+1)}^{\text{w}(ms),a} = P_{(t+1)}^{\text{w}(legal),a} - P_{(t+1)}^{\text{w}(cl),a}$$

where

Equation 9.14:

$$P_{(t+1)}^{\text{nm}(cl)} = P_{(t+1)}^{\text{nm}(legal)} \times \frac{P_c^{\text{nm}(cl)}}{P_c^{\text{nm}(legal)}}$$

Equation 9.15:

$$P_{(t+1)}^{\text{v}(cl)} = P_{(t+1)}^{\text{v}(legal)} \times \frac{P_c^{\text{v}(cl)}}{P_c^{\text{v}(legal)}}$$

Equation 9.16:

$$P_{(t+1)}^{\text{w}(cl)} = P_{(t+1)}^{\text{w}(legal)} \times \frac{P_c^{\text{w}(cl)}}{P_c^{\text{w}(legal)}}$$

where

$P_{(t+1)}^{\text{nm}(ms),a}$ = population of single persons according to the marital status concept, by age a, at time (t+1);

$P_{(t+1)}^{\text{nm}(legal),a}$ = population of single persons according to legal marital status concept, by age a, at time (t+1);

$P_{(t+1)}^{\text{nm}(cl),a}$ = population of persons living in common-law unions, never previously married, by age a, at time (t+1);

$P_{(t+1)}^{\text{mar}(ms),a}$ = population of married persons according to the marital status concept, by age a, at time (t+1);

$P_{(t+1)}^{\text{mar}(legal),a}$ = population of married persons according to the legal marital status, by age a, at time (t+1);

$P_{(t+1)}^{\text{w}(cl),a}$ = population of persons living in common-law unions who were widowed by their previous spouse, by age a, at time (t+1);

$P_{(t+1)}^{\text{v}(cl),a}$ = population of persons living in common-law unions who divorced their previous spouse, by age a, at time (t+1);

- $P_{(t+1)}^{v(ms),a}$ = population of divorced persons according to marital status concept, by age a, at time (t+1);
- $P_{(t+1)}^{v(legal),a}$ = population of divorced persons according to legal marital status concept, by age a, at time (t+1);
- $P_{(t+1)}^{w(ms),a}$ = population of widowed persons according to the marital status concept, by age a, at time (t+1);
- $P_{(t+1)}^{w(legal),a}$ = population of widowed persons according to legal marital status concept, by age a, at time (t+1);
- $P_c^{nm(cl)}$ = censal estimate of population of single persons living in common-law unions;
- $P_c^{nm(legal)}$ = censal estimate of population of single persons according to legal marital status concept;
- $P_c^{v(cl)}$ = censal estimate of population of divorced persons living in common-law unions;
- $P_c^{v(legal)}$ = censal estimate of population of divorced persons according to legal marital status concept;
- $P_c^{w(cl)}$ = censal estimate of population of widowed persons living in common-law unions;
- $P_c^{w(legal)}$ = censal estimate of population of widowed persons according to legal marital status concept.

9.6.2 Levels of estimate

The difference between preliminary and final postcensal estimates stems from the estimates by legal marital status. The same estimation method is used in both sets of estimates. As a result, the estimates by marital status have the same revision level as the estimates by legal marital status from which they are derived.

9.7 Intercensal population estimates by marital status and legal marital status

The production of intercensal estimates by age and sex is done by distributing the error of closure across age and sex cohorts. For a description on the calculation and distribution of the error of closure, see Chapter 1.

Adjusted census distributions by age, sex and marital status/legal marital status from the two most recent censuses are used to derive intercensal estimates of population by marital status/legal marital status. The census distributions are linearly interpolated to obtain the required series of distributions. The interpolated distributions are then applied to the intercensal population estimates by age and sex to obtain estimates by age, sex and marital status/legal marital status.

Chapter 10

Census family estimates

In addition to population estimates, Statistics Canada produces annual estimates of census families at national and provincial/territorial levels. These estimates are available from 1986 onwards. Estimates of the number of families⁴⁴ can be either postcensal or intercensal. Postcensal estimates are produced using data from the most recent census adjusted for net census undercoverage (NCU) and the components of growth. Intercensal estimates are produced using counts from two consecutive censuses adjusted for NCU. The breakdown of the total number of families by family type, sex and age of parents, family size and children's age is available on the basis of data from the most recent census of population. Estimates of families by characteristics are generated for provinces only.

This chapter discusses the estimation methods used for the production of postcensal and intercensal estimates of census families for Canada, provinces and territories, and estimates by characteristics at the provincial level.

10.1 Definition and relevant concepts

For the purpose of generating estimates, the definition of family is, with one exception, similar to what is used for the Census. As defined up to and including the 2001 Census, a census family can be any of the following:

- A married couple with or without children of either or both spouses living in the same dwelling;
- A couple living common-law with or without children of either or both partners;
- A lone parent of any marital status, with at least one child living in the same dwelling.

A couple living common-law can be an opposite-sex couple or a same-sex couple. Children may be children by birth, marriage or adoption regardless of their age or marital status as long as they live in the dwelling and do not have their spouse or child living in the dwelling. Grandchildren living with their grandparent(s) but with no parents present also constitute a census family.

The exception in the definition of family resides in how same-sex couples are recognized. While in the 2001 Census, same-sex common-law partners are creating a census family, they are excluded from Demography Division's family estimates.

Census family concepts in the 2001 Census

For the 2001 Census, a number of changes were made in the concept of the census family. The changes are listed as follows:

44. Unless otherwise indicated, the term "families" means census families.

Children in a census family can have been previously married (as long as they are not currently living with a spouse or common-law partner). Previously, children in a census family had to be “never-married”.

A grandchild living in a three-generation household where the parent (middle generation) has never-married will, contrary to previous censuses, now be considered as a child in the census family of his or her parent, provided the grandchild is not living with his or her own spouse, common-law partner, or child. Traditionally, the census family usually consisted of the two older generations.

A grandchild of another household member, where a middle-generation parent is not present, will now be considered as a child in the census family of his or her grandparent, provided the grandchild is not living with his or her own spouse, common-law partner, or child. Traditionally, such a grandchild would not be considered as the member of a census family.

Two persons living in a same-sex common-law relationship, along with any of their children residing in the household, will be considered a census family. Since it is impossible to establish trends with regard to the increase of same-sex common-law couples, this group has been excluded in the estimation of census families.

10.2 Data sources

The data sources used to estimate the components of change in the number of census families are often the same as those used to estimate the components of population change. The base for family estimates is the count of families in the census, adjusted for net undercoverage. Table 10.1 shows the sources and references of data used to generate the postcensal census family estimates.

Table 10.1
Sources and references of data used to generate the postcensal estimates of census families

Component	Sources and references (if applicable)
Base population (Censal estimate)	Census counts of families (Census of Canada, Catalogue no. 95F0312XCB2001003) Data from Reverse Record Check (RRC) (<i>Coverage, 2001 Census Technical Report (Reference Products 2001 Census)</i> , Catalogue no. 92-394-XIE and <i>Coverage, 1996 Census Technical Report</i> , Catalogue no. 92-370, Statistics Canada)
Marriages	Health Statistics Division, Statistics Canada Demography Division, Statistics Canada Census data
Divorces	Health Statistics Division, Statistics Canada Demography Division, Statistics Canada
Deaths	Health Statistics Division, Statistics Canada Demography Division, Statistics Canada Census data
Immigration	Demography Division, Statistics Canada Citizenship and Immigration Canada (CIC) data
Non-permanent residents	Demography Division, Statistics Canada CIC data
Emigration Returning emigrants	Demography Division, Statistics Canada Demography Division, Statistics Canada Census data
Net temporary emigration	Demography Division, Statistics Canada
Interprovincial migration	Demography Division, Statistics Canada Data from the Canada Revenue Agency's (CRA) Canada Child Tax Benefit (CCTB) files and the Small Area and Administrative Data Division's (SAADD) T1 Family File (T1FF) Census data
Common-law unions	Demography Division, Statistics Canada Census data

10.3 Postcensal estimates of census families, Canada, provinces and territories

There are several events which have an impact on the number of families, either through the formation of new families, the dissolution of existing families, or the migration of families. Table 10.2 lists those events and provides possible effects on the number of census families. Postcensal estimates of census families are obtained much in the same manner as are population counts, i.e., they are derived using data from the most recent census, adjusted for net census undercoverage (of families) and estimates of the components of change in families since that census.

Table 10.2
Events and hypotheses used to reconstruct changes in the number of census families

Events	Hypothesized effect of events on number of census families		
	Increase	No change	Decrease
Marriages ¹	If neither spouse is a family head	If one spouse is a family head	If both spouses are family heads
Divorces, separations, and annulments ² of legal marriages and dissolution ² of common-law unions	If there is more than one child and custody is shared	If there is at least one child and one parent has custody	If the couple is childless
Deaths of married persons ¹ and of persons living in common-law unions ²		If one spouse in a husband-wife family with children dies, or if a “non-family” married person dies	If one spouse in a childless husband-wife family, or both spouses in a husband-wife family with or without children, or a married lone-parent family head dies
Deaths of widows, widowers, or divorcees ¹		If a non-family person dies	If a lone-parent family head dies
Death ² or leaving home ² of a lone-parent family child		If at least another child still lives at home	If a lone child dies or leaves home
Births ¹	If a child is born to a non-family person or to a child at home	If a child is born to a head of a lone-parent family or to a husband and wife	
Interprovincial migration ¹	If a family arrives from another province or territory		If a family leaves for another province or territory
Common-law unions ¹	If neither partner is a family head	If one partner is a family head	If both partners are family heads

1. Refers to events for which insufficient information is available.

2. Refers to events for which no statistical information is available.

Not all events presented in Table 10.2 can be included in the estimates of change. Information on marriages, divorces, deaths, immigration, non-permanent residents, emigration and interprovincial migration is either readily available, or can be derived indirectly from administrative data sources. The lack of information on separations, annulments, and on children leaving home, however, makes the estimation of factors contributing to the formation or dissolution of a family rather difficult. In addition, the unavailability of current data on common-law union formations is a concern. However, attempts are made to estimate these types of unions and their family characteristics, as described in the following sections in this chapter.

Estimates of the total number of families are produced by the component method. This method takes into account the events that result in the formation of new families, the dissolution of existing families and the migration of families.

The component method formula for estimating the number of families in the year (t+1) is as follows:

For each province/territory:

Equation 10.1:

$$F_{(t+1)} = F_t + (F_{(t,t+1)}^{\text{MAR}} + F_{(t,t+1)}^{\text{I}} + F_{(t,t+1)}^{\Delta\text{N}} + F_{(t,t+1)}^{\text{RE}}) - (F_{(t,t+1)}^{\text{E}} + F_{(t,t+1)}^{\Delta\text{TE}} + F_{(t,t+1)}^{\text{V}} + F_{(t,t+1)}^{\text{Dmar}} + F_{(t,t+1)}^{\text{Dw}} + F_{(t,t+1)}^{\text{Dv}}) + \Delta F_{(t,t+1)}^{\text{UL}} + F_{(t,t+1)}^{\Delta\text{NPR}}$$

where

- (t,t+1) = interval between the date of the previous reference period, time t, and the reference date of the estimate, time t+1;
- $F_{(t+1)}$ = estimated number of families at time t+1;
- F_t = estimated base number of families at time t (census adjusted for NCU or most recent estimates);
- $F_{(t,t+1)}^{\text{MAR}}$ = number of new families formed by marriage between time t and t+1;
- $F_{(t,t+1)}^{\text{I}}$ = number of families immigrating from another country between time t and t+1;
- $F_{(t,t+1)}^{\Delta\text{N}}$ = net interprovincial migrant families between time t and t+1;
- $F_{(t,t+1)}^{\text{RE}}$ = number of returning emigrant families between time t and t+1;
- $F_{(t,t+1)}^{\text{E}}$ = number of emigrant families between time t and t+1;
- $F_{(t,t+1)}^{\Delta\text{TE}}$ = net number of families temporarily abroad between time t and t+1;
- $F_{(t,t+1)}^{\text{V}}$ = number of families dissolved due to divorce between time t and t+1;
- $F_{(t,t+1)}^{\text{Dmar}}$ = number of families dissolved due to deaths of married persons between time t and t+1;
- $F_{(t,t+1)}^{\text{Dw}}$ = number of families dissolved due to deaths of widowed persons between time t and t+1;
- $F_{(t,t+1)}^{\text{Dv}}$ = number of families dissolved due to deaths of divorced persons between time t and t+1;
- $F_{(t,t+1)}^{\Delta\text{UL}}$ = net number of common-law unions between time t and t+1;
- $F_{(t,t+1)}^{\Delta\text{NPR}}$ = net number of NPR families between time t and t+1.

Base number of census families

The base numbers of families were derived from the quinquennial censuses from 1986 to 2001. Those family counts were adjusted as follows:

1. adjustment of families for net census undercoverage (NCU);
2. extrapolation of the growth between the two censuses to adjust the estimates of families from Census Day to July 1;
3. addition of estimates of non-permanent residents in 1986. Since 1991, non-permanent residents have been included in the census universe.

Same as population counts, the adjustment of families for NCU is important. Net census undercoverage is the difference between the number of families that should have been enumerated but was missed (undercoverage) and the number of families that was enumerated but should not have been or that was counted more than once (overcoverage).

For the 1991 and 1996 censuses, the NCU of families is based on the coverage studies. The 1986 adjustment was calculated by multiplying the 1991 ratio of the NCU of families to the total population by the revised NCU of the 1986 total population.

Since the NCU of families could not be estimated on the basis of the 2001 Census coverage studies, another method had to be used. The 2001 NCU rate for families is based on the ratios of the 1996 NCU rates for families and married persons (estimates weighted by province and territory). The ratios of the 1996 rates offered the best correlation between NCU of families and the NCU of a subpopulation for which a 2001 estimate was available. The 1996 ratios were then applied to the 2001 NCU rates for married persons to obtain the 2001 NCU rates for families for the provinces and territories.

Components of the population growth of census families

The following sections describe how each of the components in the estimation of the total number of census families is derived.

The contribution of marriages to family formation

In order to estimate the number of marriages contributing to family formation during a period, marriages must first be classified by age, sex, marital status and the headship status of both partners immediately preceding marriage. The Vital Statistics database provides data on marriages, by age and prior marital status, separately for males and females. The database, however, does not provide information on headship status. This information can be estimated from census data. The combined information on marriages (from Vital Statistics) and headship rates (derived from the last census) is used to estimate the number of newly formed families.

The first step in the process is the estimation of the number of marriages, separately for each of the possible sex and marital status combinations, where at least one spouse was the head of a lone-parent family prior to the current marriage. This is accomplished by applying the headship rate for lone-

parent families of a given marital status to the number of marriages for a given sex and marital status, both at a given age, as follows:

$$\text{Equation 10.2:} \quad \mathbf{H}^{\text{ms},a} = \frac{\mathbf{L}^{\text{ms},a}}{\mathbf{P}^{\text{ms},a}}$$

where

$\mathbf{H}^{\text{ms},a}$ = headship rate by marital status and age, according to census data;

$\mathbf{L}^{\text{ms},a}$ = number of persons who are lone-parent family head by marital status and age, enumerated in the last census;

$\mathbf{P}^{\text{ms},a}$ = total number of persons by marital status and age enumerated in the last census.

$$\text{Equation 10.3:} \quad \text{MAR}^{\text{ms},a}(h) = \text{MAR}^{\text{ms},a} \times \mathbf{H}^{\text{ms},a}$$

where

$\text{MAR}^{\text{ms},a}(h)$ = number of marriages of family heads by marital status (single, widowed or divorced) and age;

$\text{MAR}^{\text{ms},a}$ = total number of marriages by marital status and age;

The age, sex and marital status-specific number of marriages for non-family heads is obtained by subtracting the number of marriages of lone-parent heads from the corresponding total number of marriages.

$$\text{Equation 10.4:} \quad \text{MAR}^{\text{ms},a}(\bar{h}) = \text{MAR}^{\text{ms},a} - \text{MAR}^{\text{ms},a}(h)$$

where

$\text{MAR}^{\text{ms},a}(\bar{h})$ = number of marriages of persons other than family heads by marital status and age.

The next step is to estimate the number of newly-married persons of a given sex and marital status, according to the headship and marital status of his/her partner (reference spouse) prior to marriage. In order to achieve this, a joint probability of marriage according to the various marital and headship status combinations, irrespective of age, must first be calculated as follows:

$$\text{Equation 10.5:} \quad \pi^{\text{ms}}(h') = \sum_{\text{ms}'} \left(\frac{\text{MAR}^{\text{ms},\text{ms}'}}{\text{MAR}^{\text{ms}}} \times \frac{\text{MAR}^{\text{ms}'}(h')}{\text{MAR}^{\text{ms}'}} \right)$$

where

$\pi^{ms}(h')$ = probability, for each marital status of reference spouse, that his/her marital partner is a family head;

$\frac{MAR^{ms,ms'}}{MAR^{ms}}$ = probability that a male (female) of marital status ms (ms') has a married female (male) of marital status ms' (ms);

$\frac{MAR^{ms'}(h')}{MAR^{ms'}}$ = probability that a newlywed female (male) of marital status ms' (ms) is already a family head, h' (h); and

$$\text{Equation 10.6:} \quad MAR^{ms'}(h') = \sum_a (MAR^{ms',a'} \times H^{ms',a'})$$

$$\text{Equation 10.7:} \quad \pi^{ms}(\bar{h}') = 1 - \pi^{ms}(h')$$

where

ms, ms' = marital status of reference spouse and that of his/her partner, respectively;

h, h' = indicator of lone-parenthood status for reference spouse and his/her partner, respectively;

a, a' = age of reference spouse and that of his/her partner, respectively;

$MAR^{ms,ms'}$ = the number of reference spouses of a given marital status, by his/her partner's marital status;

$\pi^{ms}(\bar{h}')$ = complement to $\pi^{ms}(h')$, representing the probability, for each marital status of the reference spouse, that his/her marital partner is *not* a family head.

The above probabilities, when applied to the number of marriages classified by marital and headship status of the partner, yield the number of marriages by marital and headship status as shown below:

Category 1: male only is family head

$$\text{Equation 10.8:} \quad MAR^{ms,a}(h, \bar{h}') = MAR^{ms,a}(h) \times \pi^{ms}(\bar{h}')$$

Category 2: female only is family head

$$\text{Equation 10.9:} \quad MAR^{ms,a}(\bar{h}, h') = MAR^{ms,a}(\bar{h}') \times \pi^{ms}(h)$$

Category 3: both male and female are family head

$$\text{Equation 10.10:} \quad MAR^{ms,a}(h, h') = MAR^{ms,a}(h) \times \pi^{ms}(h')$$

Category 4: neither male nor female is family head

$$\text{Equation 10.11:} \quad MAR^{ms,a}(\bar{h}, \bar{h}') = MAR^{ms,a}(\bar{h}') \times \pi^{ms}(\bar{h}')$$

In categories 1 and 2, the marriage results neither in the formation nor the dissolution of a family. In category 3, marriage eliminates two existing families and creates a new family. In category 4, a new family is formed. Thus, the probability of net family formation through marriage by age group and marital status, by sex, is given by:

Equation 10.12:

$$\pi^{ms,a} = \frac{\text{MAR}^{ms,a}(\bar{h}, \bar{h}') - \text{MAR}^{ms,a}(h, h')}{\text{MAR}^{ms,a}}$$

The final step is the estimation of the net number of families formed through marriages $F_{(t,t+1)}^m$:

Equation 10.13:

$$F_{(t,t+1)}^{mar} = \frac{1}{2} \left[\sum_{ms} \sum_a (\pi^{ms,a} \times \text{MAR}_{(t,t+1)}^{ms,a}) + \sum_{ms'} \sum_{a'} (\pi^{ms',a'} \times \text{MAR}_{(t,t+1)}^{ms',a'}) \right]$$

The contribution of divorces to family dissolution

When a couple who has no children living at home undergoes a divorce, a family is dissolved. If there are children living at home, however, the effect of divorce on net family formation is uncertain because of the lack of cross-classified information on child custody by place of residence, as well as the number of children living at home. The estimated number of family dissolutions resulting from divorce is assumed to be equal to the number of divorces involving couples with no dependent children.⁴⁵ The dissolution of families due to divorce is estimated by multiplying the number of divorces by the proportion of divorces not involving dependent children based on data published by Health Statistics Division.

The contribution of death to family dissolution

The families at risk of being dissolved through death can be: two-parent families with or without children (i.e., death of one parent in two-parent families without children, death of both parents, deaths of parent(s) and child(ren)), and lone-parent families whose head is either single, widowed, separated or divorced (i.e., death of a parent or death of an only child). No information is available on, nor has any attempt been made to estimate family dissolution through either of the following categories: (i) death of an only child in lone-parent families; or (ii) death of either children, or both parents, in the case of two-parent families. These cases are considered rare and have been ignored here. The following describes the procedures used for estimating the dissolution of family types due to death.

45. Dependent children are defined as all children under the age of 16 years, and all children over the age of 16 who are financially dependent.

Couples without children and lone-parent families whose head is a separated person

In the current methodology, separate estimation of the number of family dissolutions due to deaths in husband-wife families⁴⁶ or in lone-parent families (where the head of the family is separated) is not performed. Estimation of the total number of such family dissolutions is achieved by multiplying the number of deaths of married persons, by age group, by age-specific probabilities of family dissolution due to death, as shown below.

Equation 10.14:

$$F_{(t,t+1)}^{D_{mar}} = \sum_a (D_{(t,t+1)}^{mar,a} \times \pi^{mar,a})$$

where

$F_{(t,t+1)}^{D_{mar}}$ = number of families dissolved through deaths of married persons during time t and t+1;

$D_{(t,t+1)}^{mar,a}$ = number of deaths of married persons by age group during time t and t+1;

$\pi^{mar,a}$ = age-specific probabilities of family dissolution due to deaths of married persons.

The age-specific probabilities of family dissolution are calculated as the ratio of the sum of the number of couples without children and the number of separated lone-parents, to the total number of married persons of the same age, in the last census. The probabilities are derived as follows:

Equation 10.15:

$$\pi^{mar,a} = \frac{C^{mar,a}(c) + L^{mar,a}}{P^{mar,a}}$$

where

$\pi^{mar,a}$ = probability of family dissolution due to deaths of married persons, by age group;

$C^{mar,a}(c)$ = number of couples without children, by age group, as enumerated in last census;

$L^{mar,a}$ = number of lone parents who are separated, by age group, as enumerated in last census;

$P^{mar,a}$ = total number of married persons, by age group, as enumerated in last census.

Dissolution of families whose head is widowed or divorced

The number of families dissolved by the death of widowed and divorced persons is calculated by applying the family headship rates (Equation 10.16) in these two marital groups, to the corresponding number of deaths. The family headship rates are calculated from the most recent census, and are, in this case, used as probabilities of family dissolution. Thus:

46. Vital Statistics data currently differentiate between deaths of separated persons from those of other married persons. The required adjustment to the estimation method has however not been developed at this stage.

Equation 10.16:
$$\mathbf{H}^{w,a} = \frac{\mathbf{L}^{w,a}}{\mathbf{W}^a} \quad \text{and} \quad \mathbf{H}^{v,a} = \frac{\mathbf{L}^{v,a}}{\mathbf{V}^a}$$

where

- $\mathbf{H}^{w,a}$ and $\mathbf{H}^{v,a}$ = headship rate of widowed population and of the divorced population, respectively, by age group, in the last census;
- $\mathbf{L}^{w,a}$ and $\mathbf{L}^{v,a}$ = number of lone-parents who were widowed and divorced, respectively, by age group, in the last census;
- \mathbf{W}^a and \mathbf{V}^a = number of widowed and divorced persons, respectively, by age group, enumerated in the last census.

Equation 10.17:
$$F_{(t,t+1)}^{D_w} = \sum_a (D_{(t,t+1)}^{w,a} \times \mathbf{H}^{w,a}) \quad \text{and}$$

Equation 10.18:
$$F_{(t,t+1)}^{D_v} = \sum_a (D_{(t,t+1)}^{v,a} \times \mathbf{H}^{v,a})$$

where

- $F_{(t,t+1)}^{D_w}$ and $F_{(t,t+1)}^{D_v}$ = number of families dissolved through deaths of widowed and divorced persons, respectively;
- $D_{(t,t+1)}^{w,a}$ and $D_{(t,t+1)}^{v,a}$ = number of deaths of widowed and of divorced persons, by age group, respectively.

The total number of families dissolved by the death of widowed and divorced persons is obtained by summing the results of equations 10.17 and 10.18.

Immigration and emigration of families

The number of new families to Canada during a reference period resulting from immigration requires information on the number of family units who immigrate, and on the number of family reunifications resulting from the arrival of persons who already have a parent, child or spouse in Canada. As neither is available, certain assumptions must be made.

For estimation purposes, it is assumed that the number of immigrant or emigrant families (including families emigrating permanently and net temporary emigration of families) is closely approximated by the total number of married immigrant or emigrant females (permanent or temporary). Therefore, for every married immigrant, permanent or temporary emigrant female, it is assumed that there is one immigrant or emigrant family. For more information on immigration and emigration estimates, please see Chapters 4 and 6, respectively.

Changes in returning emigrant families

The number of returning emigrant families is calculated using the proportion of returning emigrants living in a family and the average number of returning emigrants per family according to the census used to estimate the base number of families (one-year mobility question). This ratio is applied to the estimated number of returning emigrants.

Changes in non-permanent resident families

The estimated number of non-permanent families is based on the same assumption as for estimates of immigrating and emigrating families, that is, for each married female non-permanent resident there is one non-permanent family. The annual net numbers of families are based on the estimated number of married female non-permanent residents on July 1 of two consecutive years.

Interprovincial migration of families

The base data used to produce estimates of the interprovincial migration of families are taken from the child tax benefits file (CCTB) from Canada Revenue Agency. To compensate for the CCTB's partial coverage, the ratio of the child-migrant matrix adjusted for program incompleteness to the child-migrant matrix was applied to monthly family interprovincial migration matrices:

$$\text{Equation 10.19: } {}^m M_{adj}^{Fam} = {}^m M_{CCTB}^{Fam} \times \frac{{}^m M_{adj}^{child}}{{}^m M_{CCTB}^{child}}$$

$$\text{Equation 10.20: } {}^{j-j} M_{adj}^{Fam} = \sum_{m=July}^{June} M_{adj}^{Fam}$$

where:

- ${}^m M_{adj}^{Fam}$ = adjusted family-migrant matrix for month m;
- ${}^m M_{CCTB}^{Fam}$ = family-migrant matrix based on the CCTB for month m;
- ${}^m M_{adj}^{child}$ = adjusted child-migrant matrix for month m;
- ${}^m M_{CCTB}^{child}$ = child-migrant matrix based on the CCTB for month m;
- ${}^{j-j} M_{adj}^{Fam}$ = adjusted family interprovincial migration matrix for the period from July to June.

The estimates are also adjusted to take into account families with no children and with no children under 18. The adjustment is based on the family counts from the census used to estimate the base number of families adjusted to July 1 and on the estimates for the month of July of the same year from the CCTB.

$$\text{Equation 10.21: } {}_{(j,k)} F_{(t,t+1)} = \frac{F}{F_{(j,k)}} \times {}_{(j,k)} F_{(t,t+1)}^{0-17}$$

where

- ${}_{(j,k)}F_{(t,t+1)}$ = total number of families (with and without children) migrating from province j to province k ;
 \mathbf{F} = number of families according to census data, adjusted for net undercoverage;
 ${}^{\text{CCTB}}\mathbf{F}$ = number of CCTB-receiving families;
 ${}_{(j,k)}F_{(t,t+1)}^{0-17}$ = number of families with children aged 0 to 17 migrating from province j to province k .

Contribution of common-law families to family formation

The net number of common-law families is obtained by extrapolating the population counts from the census used to estimate the base number of families and the previous census.

For each province and territory and each five-year age group, the estimate for year t is generated by:

$$\text{Equation 10.22: } U_t = \left\{ \frac{U_{2001}}{P_{2001}} + \left[\frac{t - 2001}{5} \right] \times \left[\frac{U_{2001}}{P_{2001}} - \frac{U_{1996}}{P_{1996}} \right] \right\} \times P_t$$

$$\text{Equation 10.23: } \Delta U_{(t-1,t)} = U_t - U_{t-1}$$

where

- U_t = number of common-law unions at time t ;
 P_t = population at time t ;
 $\Delta U_{(t-1,t)}$ = net change in common-law unions between times $t-1$ and t .

10.3.1 Levels of estimates

Updating estimates of the number of census families between censuses entails the use of data from administrative files or surveys. The quality of estimates of families therefore depends on the availability of a number of administrative data files that are provided to Statistics Canada by Canadian and foreign government departments. Since some components are not available until several months after the reference date, three kinds of postcensal estimates are produced: preliminary postcensal (PP), updated postcensal (PR) and final postcensal (PD). When all the components are preliminary, the estimate of families is described as preliminary postcensal. When they are all final, the estimate is referred to as final postcensal. Any other combination of levels is referred to as updated postcensal. The time lag between the reference date and the release date is five to six months for the preliminary estimates and two to three years for the final estimates.

10.4 Intercensal estimates of the number of census families

As with population estimates, intercensal estimates of census families are produced for reference dates between censuses. However, the method used is different than the one applied for population estimates. Intercensal estimates of census families are produced by linear interpolation between the counts of families in two consecutive censuses adjusted for net census undercoverage and growth between Census Day and July 1. The postcensal estimates of census families for reference dates between these two censuses are not considered.

10.5 Characteristics of census families

The following characteristics of families are estimated using a linear extrapolation of trends based on data from the two most recent censuses⁴⁷:

- size of family (2, 3, 4, 5 and 6+ members) and total number of persons in families;
- type of family, as husband-wife or lone-parent;
- broad age group of children;
- broad age group of husband and wife;
- broad age group and sex of lone-parent.

The following four operations are involved in this procedure:

- (i) extrapolation of proportions of families by characteristics for each province;
- (ii) adjustment for extrapolated proportions less than zero;
- (iii) estimation of the number of families by characteristics;
- (iv) estimation of the number of persons in census families.

(i) Extrapolation of proportions

The proportion of the number of census families (Φ) classified by characteristic is calculated for each province, for the census years ($t - 5$) and t . From these two sets of proportions, linear extrapolation produces a table containing annual estimates of Φ for postcensal years, as shown below:

Equation 10.24:

$$\Phi_{(t+i)}^k = \Phi_t^k + \left[\frac{i}{5} (\Phi_t^k - \Phi_{(t-5)}^k) \right]$$

where

Φ^k = proportion of census families with characteristic k out of the total number of census families in a province.

47. These extrapolations are based on census data, which have not been adjusted for net undercoverage, since estimates of net undercoverage are not available for family characteristics.

A separate table is calculated for each province and time interval ($t, t + i$). The sum of all cells in a table, ($\sum \Phi_{(t+i)}^k$) is equal to one.

(ii) *Adjustment for negative proportions*

Linear extrapolation can result in projected proportions, which are less than zero. If a negative proportion in any of $\Phi_{(t+i)}^k$ tables is obtained, it is replaced by a value of zero (0) and each cell is then adjusted so that the new table sums to 1. The adjusted cell values are calculated as follows:

Equation 10.25:

$$\Phi_{(t+i)}'^k = \Phi_{(t+i)}^{nk} \times \left(\frac{1}{\sum \Phi_{(t+i)}^{nk}} \right)$$

where

$\Phi_{(t+i)}'^k$ = adjusted cell value after removing negative proportions from the table;

$\Phi_{(t+i)}^{nk}$ = unadjusted cell values after removing negative proportions (these values are zero where $\Phi_{(t+i)}^k$ is negative, and equal to $\Phi_{(t+i)}^k$ otherwise, i.e., where $\Phi_{(t+i)}^k \geq 0$).

(iii) *Estimated number of census families by characteristics*

For each province, the number of families for the year ($t + i$) with characteristic k is calculated by multiplying the provincial total number of families $F_{(t+i)}$ (estimated by the component method described in the third section of this chapter) by the adjusted proportions $\Phi_{(t+i)}'^k$, specific to the province, as follows:

Equation 10.26:

$$F_{(t+i)}^k = F_{(t+i)} \times \Phi_{(t+i)}'^k$$

(iv) *Estimated number of persons in census families*

Estimating the total number of persons in census families requires several steps, due to the existence of the open-ended family size category, '6+'. First, the proportion of all persons in census families at the time of the last two censuses (Φ_t and $\Phi_{(t-5)}$), for any province, is calculated from the appropriate census data. Taking the most recent census as the base, the proportions of all persons in census families for the estimate year ($\Phi_{(t+i)}$) again by province, is calculated by linear extrapolation as follows:

Equation 10.27:

$$\Phi_{(t+i)} = \Phi_t + \left[\frac{i}{5} (\Phi_t - \Phi_{(t-5)}) \right]$$

The extrapolated proportion is then applied to the estimated total population in the estimate year ($P_{(t+i)}$) to arrive at the estimated number of persons in all census families ($P_{(t+i)}^F$), as follows:

Equation 10.28:
$$P_{(t+i)}^F = P_{(t+i)} \times \Phi_{(t+i)}$$

The next step involves calculating the proportion of all persons in census families in the estimate year who are in families of at least 6 persons ($\Phi_{(t+i)}^{6+}$). This is again accomplished by extrapolation from the two most recent censuses (t and $t - 5$), such that:

Equation 10.29:
$$\Phi_{(t+i)}^{6+} = \Phi_t^{6+} + \left[\frac{i}{5} (\Phi_t^{6+} - \Phi_{(t-5)}^{6+}) \right]$$

Next, the estimated number of persons in census families of at least 6 persons ($P_{(t+i)}^{F,6+}$) is calculated by applying the result of Equation 10.28 to that of Equation 10.27. Specifically:

Equation 10.30:
$$P_{(t+i)}^{F,6+} = P_{(t+i)}^F \times \Phi_{(t+i)}^{6+}$$

Finally, having derived the number of families of 2 through 5 persons from Equation 10.26 and the estimate of the number of persons in families of 6 or more, the adjusted estimate of the total number of persons living in census families $P_{(t+i)}'^F$ can be obtained by using the following equation:

Equation 10.31:
$$P_{(t+i)}'^F = \left[\sum_{n=2}^5 (n(F_{(t+i)}^n)) + P_{(t+i)}^{F,6+} \right]$$

where

- $P_{(t+i)}'^F$ = total number of persons living in census families;
- n = number of persons in the family;
- $F_{(t+i)}^n$ = number of families of n persons ($n = 2, 3, 4$ or 5);
- $P_{(t+i)}^{F,6+}$ = number of persons living in families of 6 or more persons.

10.6 Intercensal estimates of census families by characteristics

The following characteristics of families are estimated for the intercensal period (for each province) using a linear interpolation of trends based on data from the two most recent censuses (census t and census $(t - 5)$):

- family type (husband-wife or lone-parent);
- family size (2, 3, 4, 5 and 6 or more members);
- age of children;
- age of parents in husband-wife families and age and sex of parent in lone-parent families.

The method used is based on the change in the selected characteristics between the two censuses. The following steps are carried out for each province:

- (i) the distributions of families is calculated from the census t and census $(t - 5)$;
- (ii) to obtain the distributions for each intercensal year, interpolation between the proportions calculated in (i) is performed;
- (iii) for each intercensal year, the interpolated proportions are multiplied by the estimated totals;
- (iv) the number of persons in census families by characteristic are calculated.

Appendix I

Glossary

Age

Age as of July 1.

Ageing (of a population)

An increase in the number of old persons as a percentage of the total population.

Average absolute error of closure

Defined as the mean of the absolute percent differences between the postcensal estimates on Census Day and the results of the census adjusted for the net census undercoverage for the same day.

Average age

The average age of a population or sub-group is the mean age of all individuals in the population or sub-group. In more general terms, the average is the quotient obtained by dividing the sum total of a set of figures by the number of figures.

Base population

The base population refers to the population at the beginning of a period used as a reference or starting point for the estimation process. For postcensal estimates, the base population is the population enumerated in the most recent census, adjusted to July 1 and corrected for net undercoverage. The base population for intercensal estimates is the postcensal-estimated population.

Birth cohort (or generation)

Unless otherwise specified, refers here to a group of persons born within the 12-month period between July 1 of a given year and June 30 of a given year.

Censal estimate

The population enumerated in the census, corrected for net undercoverage. Censal estimates adjusted to July 1, are used as the base population for postcensal population estimates, to which the components of population growth are added (or subtracted).

Census coverage

Net census undercoverage: Difference between undercoverage and overcoverage.

Overcoverage: Number of persons who should not have been counted in the census or who were counted or were counted more than once.

Undercoverage: Number of persons who were intended to have been enumerated in a census but were not enumerated.

Census division (CD)

Refers to the general term applying to geographic areas established by provincial or territorial law, which are intermediate geographic areas between the census subdivision and the province or territory (e.g., divisions, counties, regional districts, regional municipalities and seven other types of geographic areas made up of groups of census subdivisions).

In Newfoundland and Labrador, Manitoba, Saskatchewan, Alberta, Yukon Territory, Northwest Territories and Nunavut, provincial or territorial law does not provide for these administrative geographic areas. Therefore, census divisions have been created by Statistics Canada in cooperation with these provinces or territories.

Census family

Refers to a married couple (with or without children of either or both spouses), a couple living common-law (with or without children of either or both partners) or a lone parent (of any marital status), with at least one child living in the same dwelling. The “kids” in a census family include the grandchild who live in the household of at least one of their grand-parents, in the absence of the parents.

Contrary to the definition used in the 2001 Census dictionary, same sex couples living common-law are excluded from the Census family estimates.

Census family structure

Refers to the classification of census families into married couples (with or without children of either or both spouses), common-law couples (with or without children of either or both partners), and lone-parent families by sex of parent. A couple living common-law may be of opposite or same sex.

Census metropolitan area (CMA)

The general concept of a census metropolitan area (CMA) is one of a very large urban area, together with adjacent urban and rural areas which have a high degree of economic and social integration with that urban area.

A CMA is delineated around an urban area (called the urbanized core and having a population of at least 100,000 (based on the previous census). Once an area becomes a CMA, it is retained in the program even if its population subsequently declines.

Census metropolitan areas (CMAs) are comprised of one or more census subdivisions (CSDs) which meet at least one of the following criteria: (1) the CSD falls completely or partly inside the urbanized core; (2) at least 50% of the employed labour force living in the CSD works in the urbanized core; or (3) at least 25% of the employed labour force working in the CSD lives in the urbanized core.

Census subdivision (CSD)

Refers to the general term applying to municipalities (as determined by provincial legislation) or their equivalent (for example, Indian reserves, Indian settlements and unorganized territories).

In Newfoundland and Labrador, Nova Scotia, and British Columbia, the term also describes geographic areas that have been created by Statistics Canada in cooperation with the provinces as equivalents for municipalities for the dissemination of statistical data.

Children in family

Children in the family are defined as never-married sons and/or daughters (including adopted children and stepchildren), regardless of their age, living in the same dwelling as their parents. Ever-married sons and/or daughters of any age are considered members of their parents' family if they live in the same dwelling and have no spouse or children or grandchildren (see also Census family, Family structure and Lone-parent family).

Cohort

Represents a group of persons who have experienced a specific demographic event during a given year. Thus, the married cohort of 2001 consists of the number of persons who got married in 2001. In the case of births, persons born within a specified year are referred to as a birth cohort or a generation.

Cohort component approach

This is a method used to produce estimates by age and sex, whereby the population is aged from year to year, and the components of demographic change are organised according to age and sex cohorts. Data required for this method includes demographic events such as deaths, immigration, emigration, etc. that can be directly linked to persons belonging to the same birth and sex cohorts.

Common-law union

Union consisting of two people of opposite sex or of the same sex who live together as a couple, without being legally married.

Component method

A method of generating population estimates which uses the components of demographic change and a base population as the input.

Components of demographic change

Any of the classes of events generating population movement or variations. Births, deaths, migration, marriages, divorces and new widowhoods are the components responsible for variations since they alter either the total population or the age, sex and marital status distribution of the population.

Dependency ratio

The ratio of the combined population aged between 0 to 19 years old and the population aged 65 years and over to the population aged between 20 to 64 years old.

Economic region (ER)

Refers to a group of complete census divisions (with one exception in Ontario) created as a standard geographic unit for analysis of regional economic activity.

Within the province of Quebec, economic regions (“régions administratives”) are designated by law. In all other provinces or territories, economic regions are created by agreement between Statistics Canada and the provinces or territories concerned. Prince Edward Island and the three territories each consist of one economic region. In Ontario, there is one exception where the economic region boundary does not respect census division boundaries: the census division of Halton is split between the ER of Hamilton – Niagara Peninsula and the ER of Toronto.

Emigrant

Canadian citizen or landed immigrant who has left Canada to establish a permanent residence in another country.

Emigration

Represents departures from Canada of Canadian citizens or landed immigrants to another country, involving a change in usual place of residence. Emigration may be either temporary or permanent. Where the term ‘emigration’ is used alone in this manual, it references permanent emigration which involves severing residential ties with Canada and acquiring permanent residency in another country. (See also ‘Persons Living Temporarily Abroad’).

Empirical Bayes model

A small area modelling procedure used to reduce the overall variance of estimates which initially suffer from high sampling errors. This procedure is based on a fundamental understanding in statistics, that it is possible to reduce the variance of small domain estimates by borrowing strength from related domains. Details of the Bayes model for estimating provincial census undercoverage are found in Dick and You (1997a; 1997b).

Enumerated population

The population of an area according to an official census.

Error of closure

Difference between the postcensal estimate for the same date as the census and the results of the census adjusted for net census undercoverage.

Ever-married

A person is considered ever-married if he/she is currently married, separated, widowed or divorced.

Family structure

Refers to the classification of a census family into two groups: husband-wife family and lone-parent family.

Flows

In migration statistics, refers to the movement of persons from one geographical area to another (often involves a change in residence from one region to another). Some flows result in an increase in population for a specific area (e.g. immigration, return emigration, inflows of non-permanent residents and internal migrants), while others result in a decrease in population (e.g. emigration, outflows of non-permanent residents and internal migrants).

Generation

Unless otherwise specified, refers here to a group of persons born within a given period. The 2001 generation represents people born during the year 2001.

Immigrant

Within the framework of this publication, the term “immigrant” refers to landed immigrant or permanent resident. A permanent resident is a person who does not have Canadian citizenship at birth but was granted the right by immigration authorities to live in Canada on a permanent basis. The number of children born in other countries to parents who are Canadian citizens that reside temporarily in another country are not included in the category.

Immigration

Immigration represents entries into Canada of landed immigrants or permanent residents from another country, involving a change in usual place of residence.

Intercensal population estimate

Population estimate for intercensal years derived by using postcensal estimates and population counts from the last census.

Internal migration

Movements of persons within Canada's geographical boundaries, involving a change in usual place of residence. Internal migration denotes movements from one province or territory to another (i.e., interprovincial migration) or movements from some other smaller defined geographical unit to another (i.e., subprovincial or interprovincial migration).

International migration

International migration represents movements of population between Canada and a foreign country which involves a change in the usual place of residence. A distinction is made with regard to immigrants, emigrants, returning emigrants, net temporary emigration and net non-permanent residents.

Interprovincial migration

Interprovincial migration represents movement from one province or territory involving a permanent change of residence. A person who takes up residence in another province or territory is an out-migrant with reference to the province or territory of origin and an in-migrant with reference to the province or territory of destination.

Infra-provincial migration

(Also known as Sub-provincial migration or Intra-provincial migration) represents movement from one region to another within the same province or territory involving a permanent change of residence. A person who takes up residence in another region is an out-migrant with reference to the region of origin and an in-migrant with reference to the region of destination.

Landed immigrant

(Also known as Permanent Resident) A person who is not a Canadian citizen by birth and who has been granted the right to live in Canada permanently by immigration authorities. This person falls under one of three classes (independent, business and family), as identified and described in Canada's Immigration Act. (See also immigrant and Permanent Resident)

Legal marital status

Indicates the legal conjugal arrangement of a person. Estimates are presented in the following categories: single, legally married (including persons who are separated), widowed or divorced.

Single: Persons who have never been married, or persons whose marriage has been annulled and have not remarried.

Legally married: Persons whose spouse is living, unless a divorce has been obtained. Persons separated are also included in this category.

Widowed: Persons who have lost their spouse through death and who have not remarried.

Divorced: Persons who have obtained a divorce and who have not remarried.

Lone-parent family

A lone-parent family consists of a father or mother, with one or more never-married children living in the same dwelling or with a child who is either married, divorced or widowed as long as the child is not living with a spouse or a child..

Marital status

Refers to the conjugal status of a person, that being single, married, divorced or widowed. In demographic estimates, a distinction is made between “legal” marital status and “de facto” marital status. See Chapter 9 for details of this distinction.

Indicates the conjugal arrangement of a person. Estimates are presented in the following categories: single, married (including persons living common-law and persons who are separated), widowed or divorced.

Single: Persons who have never been married, or persons whose marriage has been annulled and who have not remarried, and do not live common-law.

Married (Including persons living common-law and persons who are separated): Persons whose spouse is living, unless a divorce has been obtained.

Widowed: Persons who have lost their spouse through death and who have not remarried, and who do not live common-law.

Divorced: Persons who have obtained a divorce and who have not remarried, and who do not live common-law.

Median age

The median age is an age "x", such that exactly one half of the population is older than "x" and the other half is younger than "x".

Natural increase

Variation in population size over a given period as a result of the difference between the numbers of births and deaths.

Net internal migration

Sum of net intraprovincial and net interprovincial migration.

Net international migration

Net international migration is obtained according to the following formula: Immigrants + returning emigrants + net non-permanent residents – (emigrants + net temporary emigration).

Net interprovincial migration

Net interprovincial migration represents the difference between in-migrants and out-migrants for a given province or territory.

Net intraprovincial migration

Net intraprovincial migration represents the difference between in-migrants and out-migrants in a given region. A region can be defined as a census division, an economic region or a census metropolitan area.

Net non-permanent residents

Net non-permanent residents represent the variation in the number of non-permanent residents between two dates.

Never-married

A person who is not now, and never has been married.

Non-permanent residents

Non-permanent residents are persons who are lawfully in Canada on a temporary basis under the authority of a temporary resident permit, along with members of their family living with them. Non-permanent residents include foreign workers, foreign students, the humanitarian population and other temporary residents. The humanitarian population includes refugee claimants and temporary residents who are allowed to remain in Canada on humanitarian grounds and are not categorized as either foreign workers or foreign students. For Demography Division, the terms non-permanent resident and temporary resident are equivalent.

Children born in Canada to persons of non-permanent resident status are not considered non-permanent residents, but Canadian citizens.

Net temporary emigration

Net temporary emigration represents the variation in the number of temporary emigrants between two dates. Temporary emigration includes Canadian citizens and immigrants living temporarily abroad who have not maintained a usual place of residence in Canada. The estimates of persons living temporarily abroad are established base on the reverse record check (RRC). Emigrants' remains temporary if the person living abroad has intention of returning to Canada in more than six months. If the temporary emigrant's returns is unknown and the living temporarily abroad is between 6 to 24 months we consider the person as living abroad temporarily.

Permanent resident

(Also known as Landed Immigrant) A person who is not a Canadian citizen by birth and who has been granted the right to live in Canada permanently by immigration authorities. This person falls under one of three classes (independent, business and family), as identified and described in Canada's Immigration Act. (See also immigrant and Landed Immigrant).

Persons living temporarily abroad

Canadian citizens and landed immigrants who have migrated from Canada to another country for the purpose of settling there temporarily (i.e., temporary emigration).

Population

Estimated population and population according to the census are both defined as being the number of Canadians whose usual place of residence is in that area, regardless of where they happened to be on Census Day. Also included are any Canadians staying in a dwelling in that area on Census Day and having no usual place of residence elsewhere in Canada, as well as those considered non-permanent residents.

Population estimate

Postcensal: Population estimate produced by using data from the most recent available census adjusted for net census undercoverage and estimate of the population growth since that last census. This estimate can be preliminary, updated or final.

Intercensal: Population estimate derived by using postcensal estimates and data adjusted for net census undercoverage of censuses preceding and following the year in question.

Population growth or total growth

Variation of population size between two dates. It can also be obtained by summing the natural increase, total net migration and if possible, residual deviation. It can be positive or negative.

Population projections

As distinct from population estimates, a projection is an attempt to quantify what the size, spatial distribution and/or demographic characteristics of a population will be in the future, based on assumptions about future demographic trends. Population estimates are used as the base population for projections.

Precocity error

Difference between preliminary and final estimate of a particular component in terms of its relative proportion of the total population for the relevant geographical area. It can be calculated for both population estimates and for each component of population growth.

Rate

Refers to the ratio of the number of events estimated in a period (from time period t to $t+1$, usually a year is the period) to the average populations at the beginning and the end of the period. In this regard, the rates of births, deaths, divorces, marriages, etc are calculated. Generally, the rate is expressed in per 1,000.

Demographic growth or population growth: Ratio of population growth between the year t and $t+1$, to the average population of both these years. The rate is generally expressed in per 1,000.

Net census undercoverage of population: Difference between undercoverage rate and overcoverage rate.

Overcoverage of population: The ratio of the number of persons who should not have been counted in the census or who were counted more than once to the total number of persons that should have been enumerated in the census. Generally, the rate is expressed in percentage.

Undercoverage of population: The ratio of the estimated number of persons not enumerated in the census that were intended to have been enumerated to the total number of persons that should have been enumerated in the census. Generally, the rate is expressed in percentage.

Residual deviation

Difference between demographic population growth calculated using intercensal estimates of population between two dates and that obtained by the sum of the components for the same period. This deviation results from the distribution of the error of closure between years within the quinquennial period.

Returning emigrant

Canadian citizen or landed immigrant having previously emigrated from Canada and subsequently returned to the country.

Reverse Record Check (RRC)

The Reverse Record Check is one of four studies which provide estimates of census coverage error, including independent estimates of census undercoverage and overcoverage. The RRC begins by working with a sample of all persons who were enumerated (or missed) in the previous census, along with all persons who were either born or entered into Canada over the intercensal period. With the exception of a very small sub-population of returning emigrants, the RRC sampling frame includes all persons who could potentially be part of the census target universe. For a detailed description of the RRC study, please refer to Chapter 2 on Base Populations.

Sex ratio

The ratio of the number of men to the number of women. This is not to be confused with the sex ratio at birth, which is the ratio of the number of live-born boys to the number of live-born girls. This ratio is usually expressed as an index, with the number of females taken to be a base of 100.

Sprague coefficients

Series of factors which, when multiplied to a population distributed by multiples age groups, give a distribution of the same population by single years of age.

Sprague's multipliers

Sprague's multipliers are interpolation coefficients used to subdivide data. For a detailed description of the use of Sprague's multipliers, see Shryock et al. (1976).

Subprovincial migration

Also known as Infra-provincial Migration, is the movement between geographical areas within a province or territory involving a change in usual place of residence.

Total emigration

The total emigration component reflects the net flows of permanent and returning, and temporary emigrants.

Total net migration

Sum of net international and net internal migration.

Vital statistics

Includes all the demographic events (that is to say births, deaths, marriages and divorces) for which there exists a legal requirement to inform the Provincial or Territorial Registrar's Office.

Variance

The sum of the squared deviation of each observation from the arithmetic mean of all observations.

Vital events

Any event such as birth, death, marriage, divorce for which there exists a legal requirement to file a notification with the Provincial or Territorial Registrar's Office.

Year

Unless otherwise specified, the term "year" refers to the period beginning July 1 of a given year and ending June 30 of the following year.

Whittaker-Henderson graduation

A graduation technique used to obtain a smooth and consistent set of data from an irregular set of observed data.

Appendix II Symbols used in equations and their meaning

Population and demographic events (might be used for characteristics)

B	=	number of births
D	=	number of deaths
Div	=	number of divorces
E	=	number of emigrants
F'	=	number of husband-wife families
I	=	number of immigrants
IP	=	intercensal population estimate
Mar	=	number of marriages
N	=	number of interprovincial migrants
NPR	=	number of non-permanent residents
P	=	estimate of population
P	=	base population (census adjusted for net census undercoverage)
PE	=	permanent emigrants
RE	=	returning emigrants
TE	=	temporary emigration
Ve _u	=	new widowhoods

Base population (Censal estimates)

C	=	census count
\hat{D}	=	estimate of the number of persons not enumerated at their usual place of residence, as obtained from the RRC
DCS	=	number of persons included by the DCS
EN	=	number of persons who were enumerated at least one (excluding persons added by the DCS)
\hat{F}	=	adjustment factor produced by Empirical Bayes model
\hat{M}	=	net number of persons missed in single year
NU	=	net coverage error
O	=	number of persons enumerated more than once
\hat{O}	=	estimate of the number of persons enumerated more than once or in error
R	=	rate
\hat{R}	=	estimated rate
TU	=	number of persons missed in the census whether included or not by the DCS)
U	=	number of persons missed in the census and not included by the DCS)
<i>U</i>	=	undercoverage
<i>O</i>	=	overcoverage
<i>NU</i>	=	net undercoverage

Emigrants, net temporary emigrants, returning emigrants and interprovincial migration

AvgRateof_IM	=	average adult migration rate
CCTB_EM	=	number of CCTB/Tax children moving

CCTB_IM	=	migration of children according to CCTB data
\bar{F}	=	average adult estimation factor
G	=	adjustment factor for emigration propensity of non-CCTB/Tax children
NTE	=	net temporary emigrants
RateofCCTB_Tax_EM	=	emigration rate of CCTB-registered children
RateofCCTB_Tax_IM	=	out-migration rate of CCTB-registered children
RateofTax_EM	=	emigration rate of all children
RateofTax_IM	=	out-migration rate of all children
RatioAC	=	ratio of adult emigration rate over children emigration rate
τ^{RE}	=	adult to child ratio of returning emigrant population

Census family

C	=	number of couples without children in last census
H	=	headship rate
h	=	family head
\bar{h}	=	persons other than family heads
h'	=	marital partner of the reference spouse
ΔF^{ul}	=	net number of common-law unions
F	=	estimated number of families
F	=	number of families according to census data, adjusted for net undercoverage
$F^{\Delta N}$	=	net interprovincial migrant families
$F^{\Delta NPR}$	=	net number of NPR families
$F^{\Delta TE}$	=	net number of families temporarily abroad
F^{Dmar}	=	number of families dissolved due to deaths of married persons
F^{Dv}	=	number of families dissolved due to deaths of divorced persons
F^{Dw}	=	number of families dissolved due to deaths of widowed persons
F^E	=	number of emigrant families
F^I	=	number of immigrant families
F^{mar}	=	number of new families formed by marriage
F^{RE}	=	number of returning emigrant families
L	=	number of persons who are lone-parent family head
M^{child}	=	child-migrant matrix
M^{Fam}	=	family-migrant matrix
<i>n</i>	=	number of persons in the family
P^F	=	number of persons living in census families
$P^{>F}$	=	total number of persons living in census families
U	=	number of common-law unions
V	=	number of divorced persons
W	=	number of widowed persons

Characteristics (right-hand superscript)

<i>a</i>	=	age
<i>a^F</i>	=	age of females
<i>a^M</i>	=	age of mother
<i>cl</i>	=	common-law union
<i>inter</i>	=	interprovincial
<i>intra</i>	=	intraprovincial

<i>legal</i>	=	legal marital status
<i>k</i>	=	families' characteristics
<i>k_a</i>	=	age group
<i>mar</i>	=	legally married
<i>ms</i>	=	marital status
<i>ms'</i>	=	partner's marital status
<i>nm</i>	=	never married
<i>s</i>	=	sex
<i>v</i>	=	divorced
<i>w</i>	=	widowed

Time dimensions (right-hand subscript)

<i>t</i>	=	date of reference
<i>m</i>	=	month

For migration:

<i>j</i>	=	province of origin
<i>k</i>	=	province of destination

Data sources (left-hand superscript)

<i>adj</i>	=	adjusted data
<i>AMS</i>	=	automated Match Study
<i>c</i>	=	censal estimate
<i>C</i>	=	census counts
<i>CCTB</i>	=	data from the Canada Child Tax Benefit program (CCTB)
<i>CCTB_Tax</i>	=	data from the CCTB according to tax files
<i>CDS</i>	=	collective Dwelling Study
<i>Dem</i>	=	data from Demography Division
<i>f</i>	=	final estimate
<i>HS</i>	=	data from the <i>U.S. Department of Homeland Security</i> (United States)
<i>RRC</i>	=	reverse Record Check
<i>Tax</i>	=	data from tax files

Geographic area (left-hand subscript)

<i>j</i>	=	province or territory
----------	---	-----------------------

Others

Δ	=	net variation
ε	=	error of closure
Φ	=	proportion
π	=	probability
\mathfrak{R}	=	coverage rate

References

- Bédard, M. and M. Michalowski. 1997. **Advantages of the one year mobility variable for breaking down interprovincial migration by age, sex and marital status (Demographic documents No. 7)**. Ottawa: Statistics Canada, Catalogue no. 91-F0015MIE.
- Dick, P. 1995. "Modeling Net Undercoverage in the 1991 Canadian Census" in **Survey Methodology**, Vol. 21 No. 1, pp. 45-54. Ottawa: Statistics Canada.
- Fay, R.E. and R.A. Herriot. 1979. "Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data" in **Journal of the American Statistical Association**. Vol. 81, pp. 269-277.
- Gambino, J. and Dick, P. 2000. "Small Area Estimation Practices at Statistics Canada", in **Statistics in Transition: Journal of the Polish Statistical Association**. Vol. 4, pp. 597-610.
- Michalowski, M. 1999. "Canadians Residing Temporarily Abroad: Numbers, Characteristics and Estimation Methods". Statistics Canada, Demography Division, Development and Demographic Estimates Section.
- McKie, C. 1994. "Temporary Residents of Canada" in **Canadian Social Trends**, Spring 1994, pp. 12-15. Ottawa: Statistics Canada, Catalogue no. 11-008E.
- Ramsay, T. 2000. "Small Domain Estimation Based on Nonparametric Regression" (working paper). Ottawa: Statistics Canada,
- Reid, J. 1998. "Description of the Methodology Used to Create Migration Data From Taxation Records" (working paper). Ottawa: Statistics Canada, Small area and Administrative Data Division.
- Shryock, H.S., J.S. Siegel and Associates. 1976. **Studies in Population: The Methods and Materials of Demography**. Washington: Academic Press, Inc.
- Statistics Canada. 2005. **Migration Estimates – User's Guide**. Ottawa: Statistics Canada, Small Area and Administrative Data Division.
- Statistics Canada. 2004. **Coverage, 2001 Census Technical Report**. Catalogue no. 92-394-XIE. Ottawa: Minister of Industry.
- Statistics Canada. 2003a. **2001 Census Dictionary**. Catalogue no. 92-378-XIE. Ottawa: Minister of Industry.
- Statistics Canada. 2003b. **Population and Family Estimation Methods at Statistics Canada**. Catalogue no. 91-528-XIE. Ottawa: Ministry of Industry.
- Statistics Canada. 2002a. **Population and Dwelling Counts (On-line) – Highlight Tables, 2001 Census**. Catalogue no. 93F0051XIE. Ottawa: Ministry of Industry.

- Statistics Canada. 2002b. **Profile of Marital Status, Common-law Status, Families, Dwellings and Households, for Canada, Provinces, Territories, Census Divisions and Census Subdivisions, 2001 Census.** Catalogue no. 95F0487XCB2001001. Ottawa: Ministry of Industry.
- Statistics Canada. 1999a. **1996 Census Dictionary – Final Edition, Reference.** Catalogue No. 92-351-UIE. Ottawa: Minister of Industry.
- Statistics Canada. 1999b. **Coverage: 1996 Census Technical Report.** Catalogue no. 92-370-XIE. Ottawa: Ministry of Industry.
- Statistics Canada. 1987. **Population Estimation Methods, Canada.** Catalogue no. 91-528E. Ottawa: Minister of Supply and Services Canada.
- United Nations. 1998. **Recommendations on Statistics of International Migration – Revision 1.** New York. Department of Economic and Social Affairs, Statistics Division. Statistical Papers, Series M, No. 58. Rev. 1.
- Wilkinson, P. (2004). **Estimates of Internal Migration Based on New and Old Methods for Combined Annual Periods 1996-1997 to 2000-2001 (Demographic documents No. 6).** Ottawa: Statistics Canada, Catalogue no. 91-F0015MIE.