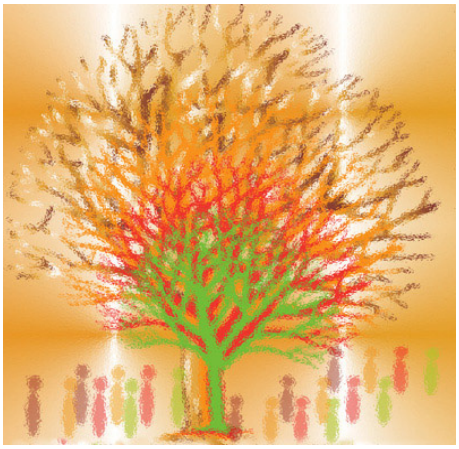
Catalogue no. 91-528-X
ISBN 978-1-100-23160-0

Population and Family Estimation Methods at Statistics Canada



by Demography Division

Release date: March 3, 2016

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

email at STATCAN.infostats-infostats.STATCAN@canada.ca

telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following toll-free numbers:

- Statistical Information Service 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
- Fax line 1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. 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.gc.ca under "Contact us" > "Standards of service to the public."

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 co-operation and goodwill.

Standard table symbols

The following 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
* significantly different from reference category ($p < 0.05$)

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2016

All rights reserved. Use of this publication is governed by the Statistics Canada Open Licence Agreement.

An HTML version is also available.

Cette publication est aussi disponible en français.

Table of Contents

List of Tables

List of Figures

Introduction

Demographic estimates are the cornerstone of statistical measurement for the population, and are put to a wide variety of uses to gain a better understanding of the economic and social situation in Canada. Their importance is highlighted by the statutory requirements for the calculation of revenue transfers and cost-sharing programs between the various levels of government. Statistics Canada, the national statistical agency responsible for this, and other data, considers the timeliness, accuracy and reliability of data to be of the utmost importance.

Every five years, Statistics Canada conducts a national census of population which collects a wide range of demographic data on the Canadian population. The census, or the National Household Survey, collects information on citizens, immigrants and non-permanent residents (NPR) residing in Canada. While some countries have a system of continuous population registration, from which basic demographic data on the state and movement of the population for non-census years can be obtained, this is not the case in Canada.

With its five-year intervals, the census cannot meet the need for timely data required by Statistics Canada's various statistical programs and external data users. To fill this gap, Statistics Canada has developed the Demographic Estimates Program. Various methodological techniques use the most recent census data collected by Statistics Canada, along with administrative data provided by other government departments and organizations, to produce estimates of the Canadian population between censuses.[1]

In 1987, Statistics Canada published its first manual detailing the methods used for estimating population, *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 agency's statistical information be open to public scrutiny. The manual was updated in 2003, 2007, 2012, and, again, herein to document conceptual and methodological changes that have been made to the Demographic Estimates Program.

Data quality indicators and analysis of demographic estimates are important for users of these estimates, and are helpful to Statistics Canada in fulfilling its role as the national statistical agency responsible for this data. Quality indicators are available in annual and quarterly publications.[2] In addition, the 1987 and 2003 versions of the report included a discussion on the data quality of the various components of demographic growth. In the previous and current versions of the report, the discussion on the quality indicators has been removed. The analysis of these quality indicators and various special data quality evaluations will be featured on their own in a companion publication on data quality indicators and analysis.

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 census net undercoverage (CNU)[3]) and estimates of the components of demographic growth since that last census. These components include births, deaths, immigration, net non-permanent residents, emigration, net temporary emigration, returning emigrants and interprovincial migration. Two more components are relevant, the residual deviation for intercensal estimates and intraprovincial migration for estimates within subprovincial areas in Canada. All elements of the production of population estimates are discussed in this manual.

Three types of postcensal estimates are produced—preliminary, updated and final estimates—referring to the timeframe in which they become available. Preliminary estimates (noted P) are typically available within three to four months after the reference date, while updated estimates (noted R) are usually available within one year. Final estimates (noted F) are the most accurate postcensal estimates available. They typically take two to three years to complete. The production of three types of estimates is the strategy that best satisfies the commitment of Statistics Canada to balance the timeliness and accuracy of data quality.

1. Statistics Canada has also developed a population projections program. Based on extrapolations of past trends, these projections reflect informed hypotheses on Canada's demographic future. For more information about the program, see *Population Projections for Canada, Provinces and Territories*, Catalogue no. 91-520-X.
2. See *Quarterly Demographic Estimates*, Catalogue no. 91-002-X; *Annual Demographic Estimates: Canada, Provinces and Territories*, Catalogue no. 91-215-X; and *Annual Demographic Estimates: Subprovincial Areas*, Catalogue no. 91-214-X.
3. Unless otherwise noted, the adjustment for census net undercoverage (CNU) also includes the partially enumerated Indian reserves.

Intercensal estimates are produced every five years and reconcile previous postcensal estimates with the latest census counts adjusted for census net undercoverage. They are generated as soon as census population counts and census net undercoverage 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: province and territory, census division, census metropolitan area, and economic region. Demographic estimates at custom-defined subprovincial levels (e.g., census subdivisions, health regions) are possible and can be estimated through special cost-recovery 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)[4] and final postcensal (PD). Intercensal estimates are produced using final postcensal estimates and counts from two consecutive censuses adjusted for census net undercoverage. According to the level of geography and type of estimate, different demographic characteristics of the population (including age and sex) are produced. Table 1 shows the level of detail and the reference period for which demographic estimates are disseminated.

Table 1
Availability of population estimates and components of demographic growth

Geography	Characteristics	Type of estimate*	Frequency
Canada, provinces and territories	Total population and components of demographic growth	• Preliminary postcensal • Updated postcensal • Final postcensal • Final intercensal	• Monthly (on request) • quarterly • annual estimates
	Population and components of demographic growth • age and sex	• Preliminary postcensal • Updated postcensal • Final postcensal • Final intercensal	Annual estimates
	Population • age, sex and marital status • age, sex and legal marital status	• Preliminary postcensal • Updated postcensal • Final postcensal • Final intercensal	Annual estimates
	Family • family size and structure	• Preliminary postcensal • Updated postcensal • Final postcensal • Final intercensal	Annual estimates
Census division / Census metropolitan area / economic region	Population and components of demographic growth • age and sex	• Preliminary postcensal • Updated postcensal • Final postcensal • Final intercensal	Annual estimates

* Note that the components are not postcensal or intercensal. They are preliminary, updated and final.

4. The acronym for updated postcensal estimates is PR, which refers to the Revised version of the preliminary postcensal estimates.

Demographic estimates at custom-defined subprovincial levels can be produced by applying synthetic estimation techniques—assuming the largest geographic area's distribution is valid for the other geographic regions or smaller ones—or by regression estimation techniques, assuming the known population totals, usually from the census, are still valid.

Users of population and family estimates

At Statistics Canada, demographic information is used to calibrate sampling weights of many social and household surveys and is a fundamental part of the analytical framework of most statistical programs. 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. The calculation of revenue transfers under various federal statutory programs, for example, as well as cost-sharing agreements between federal, provincial, territorial and municipal governments are highly dependent on demographic data. Statistics Canada is also 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 and Territorial Formula Financing (TFF), the Canada Health Transfer (CHT) and the Canada Social Transfer (CST), according to a per capita funding formula. The federal government distributes billions of dollars annually in federal transfers to provinces and territories using population estimates.[5] Since 2011, the estimates have also been used in the provincial readjustment of federal electoral boundaries as described in the *Fair Representation Act.*

The CHT and CST are federal transfers that support specific policy areas, such as health care, postsecondary education, social assistance and social services, early childhood development, and childcare.

The equalization and TFF programs provide unconditional transfers to the provinces and territories. Equalization allows 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. TFF provides territorial governments with funding to support public services to account for the higher cost of providing programs and services in the Canadian North.

Population is a key variable of the distribution formula used by the federal government to determine its annual expenditures. As billions of dollars depend directly upon these allocation formulas, accurate and up-to-date estimates are extremely important.

The Government of Canada relies on population data to make informed decisions on some of its most fundamental policies. For example, up-to-date information on the evolving demographic situation in Canada proves useful for Cabinet in its decisions relating to immigration policy. These data also contribute to increasing the cost-effectiveness of public policy expenditures.

Provincial and municipal governments also use demographic estimates to plan social programs and establish cost-sharing agreements on the basis of the per capita spending formula. Among the public service users of Statistics Canada's estimates are education and public health planners, and 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, in particular birth rates, death rates, school enrolment rates, unemployment rates, and life expectancy. Population serves directly as the denominator in many of these indicators. Survey researchers, whether at Statistics Canada, in academia, or in private polling agencies, must use up-to-date figures in survey planning and in the calculation of sampling weights.

The private sector also uses demographic estimates for business planning, marketing research and investments. The estimates also help companies properly segment and target their markets, and make sound investment decisions.

5. See "Federal Support to Provinces and Territories" on the Department of Finance Canada's website http://www.fin.gc.ca.

Overview of this manual

Chapter 1 presents the overall methods used to produce population estimates, and focuses on describing the various levels of estimation—postcensal and intercensal estimates. It explains the concept of base population and demographic growth factors and their components—from components that lead to a natural increase (births and deaths) and international effects like international migration (immigrants, emigrants, returning emigrants, net temporary emigration, and net non permanent residents)—using redistributive effects such as interprovincial migration.

The individual chapters that follow include discussions on data sources, concepts and methodology for each of the components. Each chapter describes how estimates are produced, as well as the type and origin of the data used in the calculations. This should give the reader a clearer picture of the strengths and weaknesses of each method.

Chapter 2 explains how the base population is estimated. Chapters 3 to 7 describe the components of demographic growth for Canada, the provinces and territories. Chapter 3 focuses on information on births and deaths. Chapters 4 to 6 present the components of international migration (immigration is presented in Chapter 4; net non-permanent residents in Chapter 5; and emigration, net temporary emigration and returning emigrants in Chapter 6). Chapter 7 discusses interprovincial migration.

Chapter 8 describes the methods used to produce subprovincial population estimates. Chapter 9 focuses on population estimates by age, sex, marital status and legal marital status. Chapter 10 describes the methods for estimating census families, economic entities and households. A glossary of main terms is also available.

Chapter 1

Postcensal and intercensal population estimates, Canada, provinces and territories

This chapter describes the methods 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 census net undercoverage (CNU)[6]) and estimates of the components of demographic growth since that census. The data is corrected from the Census Day to July 1 by taking into account the components of demographic growth between Census Day and June 30 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 CNU or the most recent estimate by adding and subtracting the components of demographic 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, such as births, deaths, and immigration, and those that have to be estimated, such as 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. Provincial and territorial *Vital Statistics Acts* (or equivalent legislation) render compulsory the registration of all live births and deaths within the province or territory. Vital statistics universe include births and deaths of all Canadians, immigrants and non-permanent residents (NPR) and exclude foreign residents.

International migration represents the movement of population (a change in the usual place of residence) between Canada and a foreign country.

6. Unless otherwise noted, the adjustment for the census net undercoverage (CNU) also includes the partially enumerated Indian reserves.

Figure 1.1
International migration flows for Canada

In the Demographic Estimates 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 refer to persons arriving in Canada for permanent residence (immigrants), Canadian citizens or immigrants returning to Canada after previously emigrated from Canada (returning emigrants), and Canadian citizens or immigrants leaving Canada to establish a permanent residence in another country (emigrants). Temporary flows refer to foreigners arriving for temporary stay in Canada and leaving after their stay ends (non-permanent residents), as well as Canadian citizens and 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 model based approach for the remaining components of international migration (emigration, returning emigrants, and net temporary emigration).

The last factor of demographic growth that is discussed is the interprovincial migration. While this factor does not affect the total population of Canada, it does affect the provincial and territorial population counts and is a significant challenge for the Demographic Estimates 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

Components	Sources
Base Population	• May 10, 2011 Census of Population adjusted for census net undercoverage (including the adjustment for incompletely enumerated Indian reserves and demographic adjustment if needed). • 2011 Census: Statistics Canada, Census of Canada, 2011, *Catalogue no. 98-310-X*. • Census net undercoverage: See *The Daily*, September 26, 2013. • Incompletely enumerated Indian reserves: See *The Daily*, September 26, 2013.
Births and deaths	• Statistics Canada, Health Statistics Division. • Statistics Canada, Demography Division, catalogue no. *91-215-X*, annual, catalogue no. 91-002-X quarterly.
Immigrants	Data based on the: • immigrant files provided by Citizenship and Immigration Canada (CIC).
Emigrants	Data produced by Demography Division using: • data from Canada Revenue Agency (CRA) Canada Child Tax Benefit files (CCTB) program; • tax data calculated using T1FF file provided by the Income Statistics Division of Statistics Canada; • data provided by the U.S. Department of Homeland Security, Office of Immigration Statistics; • data on the number of adult and children emigrants from T1FF file used for the provincial distribution of adults.
Returning emigrants	Data produced by Demography Division using: • data from Canada Revenue Agency (CRA) Canada Child Tax Benefit files (CCTB) program; • 2011 National Household Survey – question on the place of residence one year ago.
Net temporary emigration	Data produced by Demography Division using: • data from the Reverse Record Check (RRC) of the 2011 Census; • 2011 National Household Survey – question on the place of residence 5 years ago; • estimates of returning emigrants for 2006 to 2011 intercensal period.
Net non-permanent residents	Data produced by Demography Division using: • Field Operational Support System files (FOSS) from CIC. These files document the number of persons holding permits, authorizations or claiming refugee status. CIC is currently transitioning towards a new system: the Global case management system (GCMS). Demography division should be using this system starting with the March 2016 release.
Interprovincial migration	Data produced by Demography Division using: • CCTB-based adjusted migration data for children; • factors corresponding to the ratio of the migration rate of all children to the migration rate of eligible children calculated using CRA tax file data; • factors used to calculate adult migration and corresponding to the ratio of the adult to child migration rates, calculated on a three-year basis using CRA tax file data.

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:

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

where, for each province and territory:

$(t,t+i)$ = interval between times t and t+i;
$P_{(t+i)}$ = estimate of population at time t+i;
P_t = base population at time t (from the census after adjustment for CNU or most recent estimate);
B = number of births;
D = number of deaths;
I = number of immigrants;
E = number of emigrants;
ΔTE = net temporary emigration;
RE = number of returning emigrants;
ΔNPR = net non-permanent residents;
ΔN = net interprovincial migration.

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 and 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

The production of 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 federal, provincial 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)[7] 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 delay between the reference date and the release date is three months for preliminary estimates and two to three years for final estimates.

7. The acronym for updated postcensal estimates is PR as this level of estimates is the revised version of the preliminary postcensal estimates whereas the acronym for final postcensal estimates is PD due to its French term postcensitaires définitives.

1.2 Intercensal population estimates, Canada, provinces and territories

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

The production of intercensal estimates involves two main steps:

1. the calculation of the error of closure;
2. the linear distribution of the error of closure according to the number of days between the two censuses.

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 CNU). Assuming that the coverage studies that follow each enumeration are unbiased, the adjusted intercensal figures are considered exact.

More specifically, the error of closure is calculated as:

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

where

ε = error of closure;

P = postcensal population estimate;

$\mathbf{P}$ = census population counts after adjustment for CNU.

The error of closure comes from two sources: measurement errors in any of the components of demographic growth over the intercensal period and errors from the measurement of census coverage for the current and previous censuses.

The error of closure can be calculated for any disaggregated group, or for any summation of such disaggregation up to and including the total population. The disaggregation of the CNU is modeled as the sample size is not sufficient enough to give reliable disaggregated estimates.

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.

Once we have calculated the error of closure we are able to produce the intercensal population estimates for the five years between the two censuses. The intercensal estimates and the residuals are calculated for each month in the intercensal period.

To produce an intercensal estimate of the population at time t we need the following information:

1. the census dates (α et β).
2. the date of the intercensal estimate to be produced (t).
3. the error of closure at the end of the period (ε_β).
4. the postcensal estimate of the population at date t (P_t).

Intercensal estimates at time t are obtained using the following formula:

Equation 1.3: $$\mathrm{IP}_t = \mathrm{P}_t - \left(\frac{t-\alpha}{\beta-\alpha}\right)\varepsilon_\beta$$

Intercensal estimates are then rounded to the nearest integer.

The residual is calculated for each month in the intercensal period. This residual is an added component that is used to balance the adjustments made to the population for the error of closure. It is calculated as follows:

For the month containing the date of the previous census of the intercensal period under consideration (m(α), for example, May 2006):

Equation 1.4: $$\mathrm{Resid}_{\mathrm{m}(\alpha)} = \mathrm{P}_{\mathrm{m}(\alpha)+1} - \mathrm{IP}_{\mathrm{m}(\alpha)+1}$$

For the months m(t) between the two censuses m(α) and m(β) (for example, June 2006 to April 2011):

Equation 1.5: $$\mathrm{Resid}_{\mathrm{m(t)}} = \mathrm{P}_{\mathrm{m(t)}+1} - \mathrm{IP}_{\mathrm{m(t)}+1} - \sum_{\mathrm{k}=\mathrm{m}(\alpha)}^{\mathrm{m(t)}-1} \mathrm{Resid}_{\mathrm{k}}$$

For the month containing the date of the recent census of the intercensal period under consideration (m(β), for example, May 2011):

Equation 1.6: $$\mathrm{Resid}_{\mathrm{m}(\beta)} = \mathrm{P}_{\beta} - \mathrm{EC}_{\beta} - \sum_{\mathrm{k}=\mathrm{m}(\alpha)}^{\mathrm{m}(\beta)-1} \mathrm{Resid}_{\mathrm{k}}$$

where

EC = censal estimates.

The sum of all these residuals equals the error of closure.

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 population estimates and the census counts (after adjustment for CNU).[8] The method is the same as for the total population. The production of the intercensal estimates by age and sex involves three steps:

1. the calculation of the error of closure by age and sex;
2. the distribution of this error;
3. a final adjustment to ensure consistency with total population estimates calculated independently.

With the exception of ages between 0 and 4 years, and 100 years and over, 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 error of closure between censuses following specific cohorts generates intercensal estimates. Figure 1.2 shows the method for distributing the error of closure.

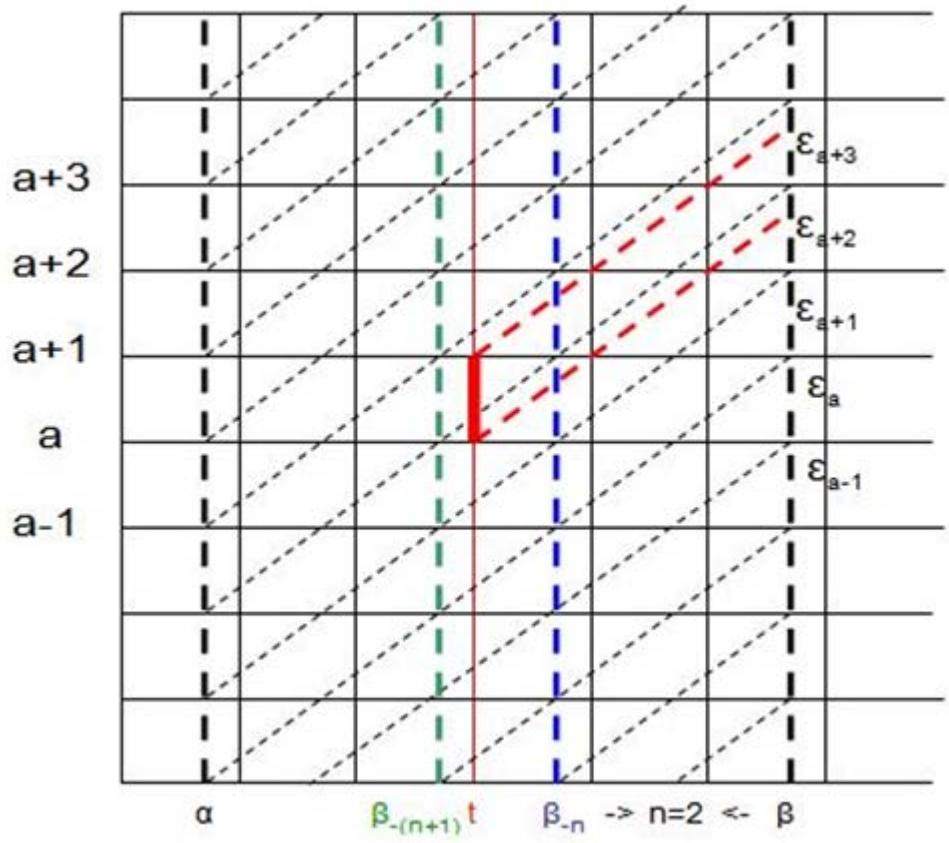
To calculate an intercensal estimate at time t for a given province (or territory) p, a particular age a and sex s, we must first define the following:

1. the dates of the two censuses (α and β).
2. the date of the estimate (t).
3. the error of closure by province, age and sex ($\varepsilon_{\mathrm{p,a,s}}$).
4. the postcensal estimates of the population at time t for province p, age a and sex s ($\mathrm{P_t(p,a,s)}$).
5. The variable n which denotes the number of whole years that separates t and β. For example, if t = 1st of July 2008 and β = 10th May 2011, then n = 2.

The following Lexis diagram (figure 1.2) is used to illustrate a general example of the intercensal estimate by age and sex.

8. The CNU by age and sex was produced by a model-based methodology as reliable estimates were not available due to insufficient sample.

Figure 1.2
Lexis diagram showing intercensal estimation



The intercensal estimate at time t for province or territory p, for age a and sex s is calculated differently based on the date and age:

A. If $t_{-a} > \alpha$ (meaning that the age cohort or part of the cohort was born after the previous census) the following formula is used:

Equation 1.7:

$$IP'_t(p,a,s) = P_t(p,a,s) - \left[\left(\frac{t-\beta_{-(n+1)}}{\beta_{-n}-\beta_{-(n+1)}}\right) f_{t,a,a+n}\varepsilon(p,a+n,s) + \left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}}\right) f_{t,a,a+n+1}\varepsilon(p,a+n+1,s)\right]$$

where

$f_{t,a,x}$ = the fraction of the age cohort at time t aged x (which is either a+n or a+n+1 at the time of the current census β), this is the portion of time between α and t in relation to the whole intercensal period.

To calculate this fraction we use:

p = date at the start of births for this cohort;

q = date at the end of births for this cohort.

These are assigned as follows:

If $x = a + n$, then $p = \beta_{-(a+n+1)}$ and $q = t_{-a}$;

If $x = a + n + 1$, then $p = t_{-(a+1)}$ and $q = \beta_{-(a+n+1)}$.

Once we have p and q, we calculate $f_{t,a,x}$ as follows:

Equation 1.8: $$f_{t,a,x} = \frac{A_{p,q}(\alpha,t)}{A_{p,q}(\alpha,\beta)}$$

where

$A_{p,q}(i,j)$ = area between time i and j of the cohort where the births have occurred between p and q.

It is noteworthy that this area is relative to the size of the cohort $(q - p)$. These results remain valid given that the size of the cohort cancels out in the calculation of $f_{t,a,x}$.

To calculate $A_{p,q}(i,j)$ we need to derive the following variables:

$i' = \max(i, p);$
$j' = \max(\min(j, q), i);$
$i'' = \max(i, q);$
$j'' = \max(j, q).$

We then calculate the area using the following formula:

Equation 1.9: $$A_{(p,q)}(i,j) = \left(\frac{(j'-i')(i'-p) + \frac{(j'-i')^2}{2}}{q-p} \right) + (j'' - i'')$$

the case where $p = q$, we set $f_{t,a,x} = 1$. This is arbitrary and will not affect the outcome since the condition to have $p = q$ implies that this term of the equation is nil.

B. If (t_{-a} <= α) and (a <= agemax - n - 2) meaning no births in the intercensal period are involved and the last age cohort is still bounded by the current census (β):

In this case the general formula described previously can still be used. In fact we can show that if (t_a <= α) and (a <= agemax - n - 2) the formula reduces to the following expression:

Equation 1.10:

$$IP'_t(p,a,s) = P_t(p,a,s) - \left[\frac{t-\alpha}{\beta-\alpha}\right] \left[\left(\frac{t-\beta_{-(n+1)}}{\beta_{-n}-\beta_{-(n+1)}} \right) \varepsilon(p,a+n,s) + \left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}} \right) \varepsilon(p,a+n+1,s) \right]$$

C. If a = agemax - n - 1; the age cohorts that will reach the last bounded age cohort during the intercensal period:

Once the age cohort reaches agemax-n-1, we have to take into account cohorts that are as old or older than the maximum age that is released in the estimates program (agemax) at the time of the recent census (β). At agemax-n-1, we use the error of closure for agemax-1 and agemax.

Equation 1.11:

$$IP'_t(p,a,s) = P_t(p,a,s) -$$

$$\left[\frac{t-\alpha}{\beta-\alpha}\right] \left[\left(\frac{t-\beta_{-(n+1)}}{\beta_{-n}-\beta_{-(n+1)}} \right) \varepsilon(p,a+n,s) + \left(\frac{\left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}} \right) P_t(p,a,s)}{\left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}} \right) P_t(p,a,s) + \sum_{i=a+1}^{agemax} P_t(p,i,s)} \right) \varepsilon(p,agemax,s) \right]$$

In the case where, $\sum_{i=a}^{agemax} P_t(p,i,s) = 0$, we suppose a uniform distribution and the equation reduces to:

Equation 1.12:

$$IP'_t(p,a,s) = P_t(p,a,s) - \left[\frac{t-\alpha}{\beta-\alpha}\right]\left[\left(\frac{t-\beta_{-(n+1)}}{\beta_{-n}-\beta_{-(n+1)}}\right)\varepsilon(p,a+n,s) + \left(\frac{\left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}}\right)}{\left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}}\right)+n+1}\right)\varepsilon(p,agemax,s)\right]$$

D. If a >= agemax - n these are the remaining cohorts in the unbounded category:

In this last case, we are looking at the age cohorts that are at agemax or older at time β.

Equation 1.13:

$$IP'_t(p,a,s) = P_t(p,a,s) - \left[\frac{t-\alpha}{\beta-\alpha}\right]\left(\frac{P_t(p,a,s)}{\left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}}\right)P_t(p,agemax-n-1,s) + \sum_{i=agemax-n}^{agemax} P_t(p,i,s)}\right)\varepsilon(p,agemax,s)$$

In the case where, $\sum_{i=agemax-n-1}^{agemax} P_t(p,i,s) = 0$, we suppose a uniform distribution and the equation reduces to:

Equation 1.14: $$IP'_t(p,a,s) = P_t(p,a,s) - \left[\frac{t-\alpha}{\beta-\alpha}\right]\left(\frac{1}{\left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}}\right)+n+1}\right)\varepsilon(p,agemax,s)$$

Adjustment of the intercensal estimate to maintain coherence with the intercensal population estimated by province

Since the error of closure is estimated by cohorts, the intercensal estimates by age and sex will not exactly match the total by province as measured in the first part of this chapter. A final adjustment is done to ensure that both estimates are consistent.

Equation 1.15: $$IP_t(p,a,s) = \left(\frac{IP_t^{total}(p)}{\sum_{i,j} IP'_t(p,i,j)}\right)IP'_t(p,a,s)$$

where, $IP_t^{total}(p)$ is the intercensal estimate measured for province p using equations 1.3 to 1.6 at time t.

The special case where the intercensal population estimates become negative

It can happen, although rarely, that for certain age and sex cohorts for certain provinces or territories that have very low counts, negative population count are assigned with the above mentioned methodology. In these cases, the counts will be set to zero and the difference will be redistributed proportionately in all the other cohorts. The estimates are then rounded to the nearest integer.

Chapter 2

Base population

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 census net undercoverage (also referred to as the censal estimate).[9] Postcensal estimates as of July 1 of a census year are calculated by the component method, using the most recent census of population adjusted for census net undercoverage (CNU)[10] and taking into account the demographic events that occurred between the Census Day and June 30. The intercensal estimates are based on postcensal estimates corrected for the error of closure.

Since census net undercoverage (census undercoverage minus census overcoverage) is an important aspect of estimating population counts used in the Demographic Estimates Program, this chapter focuses on the census population, adjusted for net census undercoverage. It begins with a brief description of the census collection procedures and the definition of the population universe of the 2011 Census, followed by a examination of the studies used to provide estimates of census coverage error with a section on adjustments for non-enumerated Indian reserves and settlements, and concludes with the procedures used for estimating census net undercoverage for the domains of age, sex and marital status.

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 has been conducted by interview. Starting from the 1971 Census, two collection methods have been used: self-enumeration and interview. In 2011, about 98% of households were enumerated using self-enumeration. A letter was sent to 60% of Canadian households. This letter replaced the traditional paper questionnaire, and provided the necessary information to enable respondents to complete the questionnaire online. Another group of households (approximately 20%) received a census package by mail. For the remaining households (approximately 20%), enumerators delivered questionnaires (approximately 18%) or completed the questionnaires during personal interviews (approximately 2%).

This interview method was normally used in remote areas of the country and on most Indian reserves. It was also used in large urban downtown areas where many residents are transient.[11]

As was done for the 2006 Census, the 2011 Census offered all households in Canada the option of completing their questionnaire online. Each letter or paper questionnaire had a unique internet access code printed on the front along with the 2011 Census website address. Respondents needed this access code to complete their questionnaire online; the information was directly submitted into the Data Processing Centre system and was verified for completeness. Approximately 54% of households responded via the Internet. Details about 2011 Census data collection and data processing procedures are described in *2011 Census Technical Report: Coverage*[12].

The base populations in the Demographic Estimates Program are derived from the quinquennial censuses between 1971 and 2011. The population universe of the 2011 Census includes the following groups:

- Canadian citizens (by birth or by naturalization) and immigrants with a usual place of residence in Canada;
- Canadian citizens (by birth or by naturalization) and immigrants who are abroad, either on a military base or attached to a diplomatic mission;
- Canadian citizens (by birth or by naturalization) and immigrants at sea or in port aboard merchant vessels under Canadian registry or Canadian government vessels;

9. The reference date for the censal estimate is the same as the corresponding census.
10. Unless otherwise noted, the adjustment for the census net undercoverage (CNU) also includes the incompletely enumerated Indian reserves.
11. The respondent was given either census or NHS questionnaire to complete and help was given if asked for.
12. See 2015d. *2011 Census Technical Report: Coverage*, Catalogue number 98-303-x, Statistics Canada, Ottawa.

- Non-permanent residents:
    - 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 study permits and members of their families living with them;
    - persons with a usual place of residence in Canada who hold work permits and members of their families living with them.

The population universe of the 2011 Census does not include foreign residents but, since 1991, non-permanent residents are included in the population universe.

Foreign residents have not been enumerated 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 definition of the population universe indicates which persons should be included in the census, but not where these persons should be enumerated. The Canadian census uses the modified *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 temporarily residents at the other location. Persons without a usual place of residence are to be enumerated wherever they happen to be on Census Day.

Each base population for the Demographic Estimates Program ($\mathbf{P}_t$, where t = the census year) is adjusted as follows (unless otherwise noted, adjustments to the base population apply to provincial, territorial and subprovincial levels):

- adjustment of the population for census net undercoverage (CNU);
- addition of independent estimates for incompletely enumerated Indian reserves in 1991, 1996, 2001, 2006 and 2011;
- adjustment for early enumeration in 1991 and 1996 in parts of northern Quebec, Newfoundland and Labrador, Yukon and the Northwest Territories;
- addition of estimates of non-permanent residents in 1971, 1976, 1981 and 1986. Since 1991, non-permanent residents are included in the census universe;
- at the provincial level, the first postcensal population estimate is July 1 of the census year. This is obtained 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 population estimate is obtained by applying to the annual components of growth, a fraction of the year that corresponds to the period between Census Day and June 30. These are adjusted to the appropriate provincial and territorial components.

2.2 Adjustment for census net undercoverage (CNU)

Coverage errors are defined as errors caused by the miscounting of the population on Census Day. There are two types of coverage error. Population undercoverage refers to the error of excluding someone who should have been enumerated. Population overcoverage refers to the error of either enumerating someone more than once or including someone who should not have been enumerated. The latter error is considered negligible. Undercoverage is more common than overcoverage. The net impact of undercoverage and overcoverage on the size of a population of interest is census net undercoverage (CNU). Census net undercoverage is calculated as the number of persons excluded who should have been enumerated (undercoverage) less the number of excess enumerations of persons enumerated more than once (overcoverage). Coverage errors are one of the most important types of error since they

affect not only the accuracy of the counts of the various census universes, but also the accuracy of all of the census data describing the characteristics of these universes.

Following each census, Statistics Canada undertakes coverage studies to measure coverage errors. Coverage studies provide undercoverage estimates for the 1991 to 2011 Censuses at the provincial and territorial levels, and for the 1971 to 1986 Censuses at the provincial level only. Estimates of overcoverage at the provincial and territorial levels are only available as of the 1991 Census. 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 CNU for Yukon and the Northwest Territories prior to 1991 was estimated by assuming that the ratio between the CNU 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 CNU 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 CNU are based on the 2011 Census coverage studies.

2.2.1 Census coverage studies

Census coverage error of the 2011 Census is measured by three studies. The 2011 Dwelling Classification Survey (DCS) addressed coverage error resulting from dwelling occupancy classification error. Census data were adjusted for this type of coverage error. The 2011 Reverse Record Check (RRC) measured population undercoverage. The 2011 Census Overcoverage Study (COS) measured population overcoverage. Census data are not adjusted for the population coverage error measured by the RRC and the COS. Rather, estimates of census net undercoverage are used in the production of Statistics Canada's demographic estimates of population.

The methodology of each of the 2011 coverage studies is described below.

A. Dwelling Classification Survey (DCS)

One of the potential sources of error in a census is the misclassification of dwellings. When a questionnaire is not returned from a household, the enumerator has to determine if the dwelling is occupied or not. Two types of errors can occur. First, an occupied dwelling can be incorrectly classified as unoccupied. This classification error results in census dwelling and population undercoverage because the dwelling is excluded from the census database. Second, an unoccupied dwelling can be incorrectly classified as occupied. When this error occurs, no questionnaire will be received for this dwelling and it will be subject to Non-response follow-up (NRFU). The dwelling will be considered as a non-respondent dwelling and therefore subject to imputation. This would add persons to the census database when, in fact, no one is living at that dwelling thus resulting in population overcoverage. Estimates from the DCS are used to adjust census data for both of these coverage errors.

An additional type of dwelling classification error measured by the DCS is the error incurred when marginal dwellings or dwellings under construction are classified in error as dwellings. This misclassification of dwellings can result in dwelling overcoverage however; census data are not adjusted for these dwellings so census estimates of the dwelling stock include some degree of overcoverage.

The DCS target population was all non-response private dwellings and all unoccupied private dwellings excluding dwellings in collective collection units (CU), canvasser CUs and Indian reserves CUs. A sample of private dwellings in the sampled CUs that were classified as unoccupied on Census Day or classified as occupied but for which no census form had been returned, was to be checked again in late June or early July 2011 to determine the true occupancy status of the dwellings on Census Day. A DCS questionnaire was used for this purpose.

At this point in processing, the unoccupied dwellings and the non-response dwellings in the sample were separated and the classification of these dwellings was confirmed against final census listing. The questionnaires completed for each sampled CU were matched to the final census listing of unoccupied dwellings. If a match could not be found, the sampled dwelling was discarded and no further processing was required. Dwellings listed as unoccupied on the census list for which no DCS questionnaire was received were considered as total non-response and went onto the next step of processing. Similarly, the final census listing of all dwellings for which a census questionnaire was not received was used to establish which of the DCS dwellings for which a DCS questionnaire was not received would be considered as total non-response.

Total non-response was addressed by a weighting adjustment while item imputation was used for item non-response. The procedure was the same for the unoccupied dwellings and non-response dwellings. When there was no information for a dwelling, the design weights of the respondents were adjusted by the design weight of the non-respondents.

Once the DCS estimates were produced, census data were adjusted for non-response dwellings and for occupied dwellings classified in error as unoccupied. This process resulted in all private dwellings on the database being classified as either occupied or unoccupied. A second procedure was used to impute the household dwelling size and other variables for the selected non-response dwelling. Household size was determined by randomly selecting a dwelling from all dwellings that had completed a census questionnaire in the same CU (nearest neighbour imputation). The complete record from this donor household was then assigned to the non-response dwelling. If no donor was found, then only a household size was assigned.

B. Reverse Record Check (RRC)

The Reverse Record Check (RRC) is a postcensal study carried out to estimate 2011 Census population undercoverage. The target population, which consisted of all persons who should have been enumerated in the 2011 Census, was formed from six sources (sampling frames). The first five frames were used to estimate undercoverage in the ten provinces, whereas estimates for the three territories were calculated based on samples from the last frame only. The six sampling frames of the 2011 RRC are:

1. 2006 Census: all persons enumerated in the 2006 Census for which names and dates of birth were completed and valid;
2. Missed: all persons from the 2006 RRC sample who were classified as missed including all persons enumerated for which names or dates of birth were missing or invalid;
3. Births: all children born between May 16, 2006 and May 9, 2011;
4. Immigrants: all immigrants who arrived in Canada between May 16, 2006 and May 9, 2011;
5. Non-permanent residents: all persons from another country, who held employment or student permits, covering May 10, 2011 and persons claiming refugee status before May 10, 2011. Family members living with them in Canada are also in this frame;
6. Health care files: all persons listed in the health care files of Yukon, the Northwest Territories, and Nunavut who were living in these territories on May 10, 2011.

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 in May 2006, and thus have been enumerable in the 2006 Census. The person would then be in both the immigrants frame and the census frame if he or she was enumerated, or in the immigrants frame and the missed frame if not enumerated. All potential cases of frame overlap must be identified to avoid double-counting.

Another difficulty is that none of the first five sampling frames covered people who had emigrated, or who were outside the country at the time of the 2006 Census without being enumerated and had returned during the intercensal period. Coverage error estimates do not include these populations.

Sampling fractions were not the same in all strata. To make the sample design more efficient, higher sampling rates were applied in subgroups for which high undercoverage or a lower tracing rate was expected.

The methodology for the territories that was changed in 2006 was once again used in 2011. As with RRCs previous to 2006, the sampling frames of the three territories were created from their respective health care files. However, the people listed in the sampling frames of each territory were then matched by name, sex and date of birth with the 2006 Census (or 2011 respectively) response database using exact matching. A manual verification was also performed. Matched people were classified as enumerated, and given a weight of 1. People not classified as enumerated were then stratified by age and sex.

After sample selection and checking the sample for quality of information for different variables of interest (i.e., geographic or demographic), the sample was ready for processing and classification. The goal of processing is to determine whether each selected person (SP) was part of the census target population and, if so, to determine whether each SP was enumerated. In addition, processing is undertaken to provide further information for the non-response adjustment.

Most of the work in processing involved searching the RRC version of the 2011 Census Response Database (RRC RDB) to determine whether the SP was enumerated at one of the addresses associated with him or her. The addresses were obtained from various sources including:

- the sampling frame for the selection address;
- updates with help from tax records;
- the computed-assisted telephone interview (CATI) and paper questionnaires;
- matches with the RRC Response Data Base (RDB) using birth date and sex of the SP and members of his or her household, or, the SP's name, postal code or telephone number as well as for all the members of his or her household.

Two outcomes could result from this process. First, when the SP was found, the classification of enumerated was usually assigned and no further processing was required. An exception was SPs who were later identified as deceased before the census from vital statistics for deaths. Second, when the SP was not found or identified as deceased before the Census, the case was sent for collection. While collection was taking place, searching the RRC RDB continued. When data from the CATI interview was available, it could be determined whether or not each SP was part of the census target population. If so, the CATI data could enable further searching.

Processing provides the information required to determine which SPs were:

1. listed;
2. mobile;
3. included in the census target population;
4. classified;
5. enumerated;
6. missed.

Selected persons for whom one or more of the above-mentioned characteristics could not be determined were considered as non-respondents. The persons that were classified were considered as partial non-responses as we knew they were part of the census target population without having enough information to determine if they were enumerated or not. Selected persons, who were in the census target population but were not enumerated, thus classified as missed, were the basis for the estimate of undercoverage.

The final weights of the selected persons (SP) began with their initial (or design) weights. The initial weight of an SP from the missed frame was the final weight assigned to him or her during the previous Reverse Record Check (RRC) when the SP was classified as missed or as being enumerated but with missing or invalid information on names or dates of birth. For the other sampling frames, the initial weights were generally equal to the inverse of the probability of selection. The exception was the non-permanent resident's frame where the initial weight was higher to account for the small number of non-permanent residents who were not in the sampling frame when the sample was selected. Final non-permanent resident counts were only available after the sample was selected. Initial weights were adjusted to add to these counts. The census frame may contain people enumerated more than once. For the first time in 2011, we set out to identify the persons selected from this frame who had been enumerated more than once in order to correct the problem. The weights of these persons were adjusted downward to compensate for the fact that they are in the frame more than once.

In order to reduce bias, the initial weights of the respondents had to be adjusted to account for non-response. The weight of the non-respondents was redistributed among the respondents. Where possible, this was done by ensuring that the weight of non-respondents with certain characteristics was redistributed only to respondents with the same characteristics. In the rare cases where a respondent with the same characteristics as a non-respondent could not be identified in a stratum, the stratum was grouped with another stratum deemed similar.

After adjusting for non-response, the estimated number of enumerated persons in the territories has traditionally been lower than the comparable census count. This is likely due to undercoverage of the census target population in the health care files. To address this bias, the weight of SPs selected in a territory was adjusted so that the estimated number of enumerated persons equaled the comparable census count for that territory.

The RRC RDB differs from the final census database in that it does not include imputations made during Whole household imputation (WHI), enumerations with an invalid or missing name or an incomplete or invalid birth date, or enumerations added after the start of the RRC data processing phase. People from the target population who are not in the RRC RDB are classified as missed. Census population undercoverage is estimated by the number (weighted) of missed persons less the number of persons excluded from the RRC RDB.

C. Census Overcoverage Study (COS)

Population overcoverage is the number of enumerations in excess of persons who are included in census tabulations more than once, usually twice. This is an error resulting in bias for census counts and estimates because they should only have been included once. Following the 2001 Census, the level of overcoverage due to duplication of individuals was measured by three studies, each one covering a part of the overcoverage: the Automated Match Study (AMS), the Collective Dwelling Study (CDS) and the Reverse Record Check (RRC). The introduction of names to the Census Response Database (RDB) since 2006 provides an opportunity to use name matching to measure overcoverage and therefore estimate overcoverage with a single study, the Census Overcoverage Study (COS). The 2011 COS is based on a series of automated exact and probabilistic matching operations and manual work. These matching operations also involve the use of various administrative data files. Therefore, since 2006 the RRC no longer measures overcoverage.

In principle, the RDB could have been matched to itself to detect duplicate enumerations. However, on a practical level, and for methodological considerations, the COS was conducted in two steps as outlined below.

Step 1 – Probabilistic matching with administrative data

The first step was based on probabilistic matching procedures, and involved matching the RDB with a set of administrative data files representing a large portion of the census target population. It was expected that this process would create a base sample of cases, including a good proportion of overcoverage cases. In particular, the majority of RDB records assigned to the same administrative record through 'many-to-one' matches were identified to be cases of overcoverage after manual review, since they pointed to the same individual from the administrative data files. A sample of these cases was chosen and verified manually to determine if they were cases of overcoverage or not and then weighted to produce the estimates.

The following administrative data files were used:

- 2005 to 2009 income tax records;
- Canadian child tax benefits;
- Birth files for Canadian citizens born between 1974 and 2008;
- Immigration files for immigrants and non-permanent residents up until September 2011;
- Health care files from Yukon, the Northwest Territories and Nunavut up until July 2011.

Since some individuals can obviously be in more than one administrative file, the files are used sequentially. As a result, once a person from the census was linked to an administrative file, no attempts were made to link that person to subsequent files. This strategy was used to prevent incidents of individuals who are in more than one file, rather than the 2006 strategy that aimed to identify all cases of overlap and remove them at the source.

Overcoverage was identified by taking a sample of cases where two or more RDB records matched to the same administrative record (or group of administrative records). The sampled cases were manually verified to determine if they were overcoverage cases. We then proceeded with the weighting and the estimation. For evaluation purposes, a sample of the one to one cases was manually verified.

In Step 1, for technical reasons, RDB records for the provinces were matched to provincial administrative records, and RDB records for the territories were matched to the records in the territorial administrative Health Care Files. Hence, cases of overcoverage between the provinces and the territories were missed at Step 1, but they were included in Step 2.

Before Step 2, the RDB was split into two parts, A and B. Part A consisted of all RDB records that were matched to at least one administrative record, whether overcovered or not. Part B consisted of all RDB records that were not

matched to an administrative record, as well as territorial records. The latter was done to take into account provincial and territorial matches that were missed in Step 1.

Step 2 – Probabilistic match with the Census Response Database (RDB)

Step 2 of the COS is a probabilistic record linkage between RDB records that were not matched with an administrative record (Part B) and the complete RDB (Part A + Part B). Statistics Canada's Generalized Record Linkage System (GRLS) was used for this step.

Within the framework of GRLS, variables such as first name, last name, sex, date of birth and some variables related to geography were considered during the record linkage. GRLS provided results in pairs of individuals with an associated weight that indicates the strength of the match. The higher the matching weight is, the more likely the pair is a good match, thus resulting in overcoverage.

The standard Fellegi-Sunter (1969) approach was implemented in the GRLS. A lower threshold, S1, was established below which matches were rejected without further review (i.e., no overcoverage), in order to minimize cases of overcoverage below threshold S1. In order to verify cases above the S1 threshold (i.e., pairs whose matching weight was greater than S1), a sample of these matches was selected for manual verification.

The household members of the persons in Steps 1 and 2 were also considered, and pairs (from the two households in question) were created using the individuals with similar data. They were then sampled and checked manually. These additional pairs were weighted and used in estimation.

In 2011, overcoverage was measured primarily by the Census Overcoverage Study (COS). The total overcoverage estimate comprised individuals overcovered in the first two steps, and those identified as overcovered among household members.

To evaluate the COS, the Automated Match Study (AMS) was repeated in 2011. The COS estimates were compared to those of the AMS. The comparison revealed a bias in the COS estimates whereby some pairs identified in the AMS were not found in the COS frames. Since the AMS provided an estimate of overcoverage not included in the COS, the last step in estimating overcoverage was to account for this bias by using the AMS estimates to adjust the COS estimates.

2.3 Calculating census net undercoverage

Let T represent the total or true number of persons in the census target population. Then, let C be the published census count of the number of persons in the census target population. The error in using C instead of T as denoted by N, and it is the census net coverage error, defined as:

Equation 2.1: $$\mathbf{N} = \mathbf{T} - \mathbf{C}$$

The censal population $\mathbf{P}$ is defined as:

Equation 2.2: $$\mathbf{P} = \mathbf{C} + \mathbf{N}$$

Let U denote population undercoverage. U is the number of persons not included in C who should have been.

Let O denote population overcoverage where O is the number of persons included in C who should not have been. There are two components to O. The first is persons who were enumerated more than once. These duplicate enumerations should not have been included in C. The census coverage studies focus on duplicate enumerations. The second component of O is persons who were included in C who are not in the census target population. Foreign residents visiting Canada, for example, who are listed on a census form as usual residents of a dwelling should not be included in C. Fictitious persons are another example. The number of persons included that are not in the census target population has been seen by previous studies to be negligibly small. Therefore, the 2006 and 2011 Census coverage studies did not measure this component of coverage error.

Since U refers to persons who should be included in C and O refers to persons who should not be included in C, the difference between T and C is U less O. That is:

Equation 2.3: $$N = U - O$$

The true number of persons in the census target population is then:

Equation 2.4: $$T = C + N = C + U - O$$

An estimate of T is given $\hat{T}$ where:

Equation 2.5: $$\hat{T} = C + \hat{N} = C + \hat{U} - \hat{O}$$

$\hat{U}$ is an estimate of the number of persons not included in C that should have been;
$\hat{O}$ is an estimate of the number of persons included in C who should not have been. Let us assume that overcoverage from persons included in C who are not in the census target population is zero. Therefore, $\hat{O}$ is restricted to an estimate of the number of duplicate enumerations. It is the goal of the census coverage studies to produce $\hat{U}$ and $\hat{O}$.

Census population coverage error can be usefully expressed as rates relative to the true population. The undercoverage rate R_U is U expressed as a percentage of T. The overcoverage rate R_O is O expressed as a percentage of T. The census net undercoverage rate R_N is the difference between U and O expressed as a percentage of the census target population. These three rates can be estimated by $\hat{R}_U$, $\hat{R}_O$ et $\hat{R}_N$ as follows:

Equation 2.6: $$\hat{R}_U = 100 \times \frac{\hat{U}}{\hat{T}} = 100 \times \frac{\hat{U}}{C+\hat{N}}$$

Equation 2.7: $$\hat{R}_O = 100 \times \frac{\hat{O}}{\hat{T}} = 100 \times \frac{\hat{O}}{C+\hat{N}}$$

Equation 2.8: $$\hat{R}_N = 100 \times \frac{\hat{N}}{\hat{T}} = 100 \times \left[\frac{\hat{U}-\hat{O}}{C+\hat{N}}\right]$$

A positive census net undercoverage rate indicates that undercoverage is larger than overcoverage. That is, there are more people not included in the published census count C than the number of duplicated enumerations. This has been, and continues to be, the experience of the Canadian census. For some domains of interest, however, negative census net undercoverage has recently been observed.

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 18 in the 2011 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. Additionally, in 2011, 13 Indian reserves in Ontario were enumerated late as they had to temporarily relocate due to forest fires.

Neither the 2011 Census nor the Reverse Record Check is in a position to produce an estimate of the population living in the 18 incompletely enumerated Indian reserves and settlements. In order to produce official estimates of population, a model-based methodology was used to prepare estimates of population for these geographical areas. For the 13 Ontario reserves, the population numbers are from a Statistics Canada enumeration that took place after the Census.

A two step model was developed to estimate the population of the 18 incompletely enumerated Indian reserves. The first step uses a simple linear regression to predict the census count in 2011. The linear regression was constructed using all Indian reserves that were completely enumerated in both the 2006 and the 2011 Census.

The model assumes a linear growth from 2006 to 2011 for all provinces with separate estimates, for the intercept and the regression parameters for each province. For each incompletely enumerated reserve, the input variable for the regression model was either the actual census count in 2006 or the best predicted census count from the 2006 model. The output of the model was the estimated census count in 2011.

The second step is done to produce consistency with the results of the census coverage studies. An adjustment was made to the estimated census count to account for census net undercoverage of all subjected census counts. Census net undercoverage for the incompletely enumerated reserves was estimated by calculating the census net undercoverage rate for all completely enumerated reserves in each province and then applying that rate to the estimated census count of all the incompletely enumerated Indian reserves in the province. The estimated census count and the estimated net missed persons in each reserve were then summed to create an estimated population for the incompletely enumerated Indian reserves. This procedure was also applied to the 13 Ontario reserves.

2.5 Net census undercoverage estimates by single year of age and sex

The Demographic Estimates Program requires an estimate of net census undercoverage for various domains of estimation. The sections above describe the methodology used in coverage studies to estimate undercoverage and overcoverage. These estimates are reliable for large domains (e.g., provinces and territories or broad age groups), but would likely be less accurate for smaller domains because of small sample sizes.

Net census undercoverage figures must be whole numbers (positive or negative), and must be produced for domains of estimation that correspond to the cross-tabulation of the following variables: single year of age, sex, provinces and territories, and subprovincial areas. They are also required for the domains composed of marital status (legal and historical), single year of age, sex, and provinces and territories.

In addition, the modelled net census undercoverage figures, when summed across domains of interest, must match the net census undercoverage totals for large domains published in the coverage studies. Furthermore, net census undercoverage must be a smooth function of single year of age; that is, the change in net census undercoverage from one single year of age to another must be regular and, ideally, must not fluctuate too rapidly or abruptly. The following sections present a detailed description of the net census undercoverage estimation methodology for small domains.

2.5.1 Provinces and territories

This section describes the methodology for estimating net census undercoverage by single year of age and sex in the provinces and territories. Normally, direct estimates (or estimates produced by the coverage studies) could be used for all domains of interest. However, direct estimates may be much too volatile as a result of the small sample size of some of these domains. As a result, a model better suited to small domains was used: the empirical Bayes or Fay-Herriot model. This model is used to produce more accurate estimates for small domains, despite their potential bias. As a result, estimating using this model is a trade-off between bias and variance in net census undercoverage estimates.

Moreover, producing estimates using the empirical Bayes model requires direct estimates and their variance for every domain of interest, as well as independent variables[13] related to net census undercoverage. In the Demographic Estimates Program, direct estimates and their variances were not available for single years of age, but were available for the intermediate domains of broad age groups (0 to 19 years, 20 to 29 years, 30 to 44 years, and 45 years and older). The totals of the independent variables for these domains were also available.

However, all the parameters of the empirical Bayes model must be estimated, specifically model error variance and the regression slope. Regression slope is estimated using the ordinary least squares technique. The independent variables are selected from a series of variables according to their contribution to R squared in the model. The selected variables ultimately result in a more parsimonious model (one that contains the fewest terms possible, while remaining highly predictive).

Model error variance was estimated using the restricted maximum likelihood technique. The results of this method were compared with the results produced using the Fay-Herriot and Wang-Fuller adjusted density maximization

13. The independent variables analyzed and the reasons for their use are presented in Dick (1995).

methods. All these methods use iterations to estimate model variance. The first method was selected since the net census undercoverage estimates that it produced generally had the smallest mean square error.

Lastly, to produce modelled estimates for the required domains (that is, by single year of age), synthetic expansion was used. It is based on the implicit assumption that the net census undercoverage adjustment factor is constant for all single years of age in an intermediate domain. For example, the net census undercoverage adjustment factor at age 1 is the same as the one applied at age 5, since both ages are in the intermediate domain of 0 to 19 years.

Formally, it is described as follows:

Equation 2.9: $$\widehat{M}_{jka} = C_{jka} \times \left(\widehat{F}_{jk} - 1\right)$$

where

$\widehat{M}_{jka}$ = net number of people omitted for a given age in province or territory j and broad age group k;

C_{jka} = number of people counted in the census for a given age a in a province or a territory j and a broad age group k;

$\widehat{F}_{jk}$ = adjustment factor produced by the empirical Bayes model by broad age group k.

As stated above, modelled net census undercoverage estimates must be consistent with published coverage study estimates. In other words, the sum of the modelled estimates must equal the net census undercoverage produced by the coverage studies in large domains, such as provinces and territories. Still, it is preferable for net census undercoverage to change gradually from year to year. To satisfy these two constraints, the modelled estimates must therefore be adequately adjusted. To this end, the raking ratio algorithm technique (Deming 1943) is applied to the modelled net census undercoverage estimates. This technique uses provincial net census undercoverage totals as the first margin and smoothed directed estimates by single year of age as the second margin.

The totals of the first margin were taken directly from the published coverage studies and were used without modification. The totals of the second margin were produced by smoothing the direct estimates by single year of age at the national level. They were then calibrated to match the total of the provincial estimates from the coverage studies.

Specifically, smoothing can be modelled using an expression such as

Equation 2.10: $$M_a = g(a) + \epsilon_a$$

where M_a is the national net census undercoverage at age a, such that a varies from 0 to 100, and g(a) is a smooth function of age.

The smoothed estimates are produced by solving the following optimization problem with constraints:

Equation 2.11: $$S(g) = \sum_{a=0}^{100} w_a (M_a - g(a))^2 + \lambda \int g''(x)\,dx$$

where w_a is a given constant, often the inverse of the variance of M_a and λ is a smoothing parameter. The smaller the value of the constant w_a the less weight is put on the direct estimate in question (which occurs when variance of the direct estimate is large). Moreover, the smoothing parameter is estimated mechanically using the generalized cross-validation criterion.

The smoothed net census undercoverage estimates are smoother than direct estimates. However, the sum of these smoothed estimates was inconsistent with the national net census undercoverage estimate published in the coverage studies. To address this inconsistency, the smoothed estimates were calibrated with the direct estimates for the broad age groups (intermediate domains) using the Denton method (1971), which is based on the principle of fluctuation preservation. In other words, the smoothed estimates are adjusted, but the year-to-year fluctuations between net census undercoverage levels are preserved. This is in line with the assumption that net census undercoverage is a smooth function of age.

2.5.2 Subprovincial areas

Base populations for census metropolitan areas and census divisions are obtained by applying the corresponding provincial and territorial census net undercoverage rates, available by age and sex. This synthetic estimate assumes that within a province or a territory and for a single year of age there is a constant census net undercoverage rate. For example, in British Colombia 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 or territorial estimates to the appropriate geographic regions. All figures sum to provincial, territorial and national totals.

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

2.6.1 Provinces and territories

At this point, estimates of census net undercoverage are available by single year of age and sex for each province and territory. However, the Demographic Estimates Program also requires estimates of provincial and territorial census net undercoverage by age, sex and marital status.

The estimates of census net undercoverage (CNU) by single year of age, sex and marital status[14] and legal marital status[15] for each province or territory were modeled by Social Survey Methods Division of Statistics Canada. The method used to estimate the marital status and legal marital status CNU is a two step raking ratio procedure. As was the case with CNU by single year of age, sex, province and territory, two margins are necessary. One margin was the previously released estimate of net missed persons by single year of age and sex for each province or territory. Whereas the other margin used was from the coverage studies showing the estimated net missed persons by age for marital status or legal marital status for each province or territory. Some minor modifications were brought to the direct survey estimates in order to ensure coherence between the estimates for marital status and the estimates modeled by single year.

The raking ratio procedure was used in a two step process. The first step was to use the five-year age groups and the second step was to use these results to produce the single year estimates. The choice was to create a design matrix using the census distribution of marital status, for each sex, within each five-year age grouping. The resulting distribution of marital status within an age group determined the structure of the design matrix. If any single marital status estimate in any five-year age group constituted at least 1 percent of the total for that age group, then the design matrix was set to one, otherwise it was set at zero. These conditions imply that the various categories are independent of each other if they meet this threshold. Finally, for each province or territory and sex, the legal marital status used the identical design matrix derived from the marital status. After convergence, the estimated CNU rates were checked, and if any estimate had a CNU rate exceeding 50 percent then the initial value for that cell was re-set to zero and the raking procedure was repeated.

Single year of age estimates by province and territory were calculated by a second raking ratio procedure. Again two marginal totals had to be fixed. One margin was the counts of previously released single year of age by sex while the other margin was the marital status or legal marital status estimates for each sex by province or territory. The initial counts were the five-year age estimates. The final single year estimates were then checked to ensure that for every single age the explicit constraint between marital status and legal marital status was maintained. This constraint means that for the adjusted population, the difference between the marital status married or common law and legal marital status married is greater than or equal to zero. On the other hand, the difference between the marital status and legal marital status for the other category (divorced, separated, single and widowed) is less than or equal to zero. The estimates were then rounded to the nearest integer all while respecting the totals already published in the coverage studies.

14. The Demography Division disseminates estimates by marital status and by legal marital status. In the estimates by marital status, persons living in a common-law union are grouped in a distinct category.

15. Legal marital status is the legal marital status of a person. It encompasses the five following categories: married and (not separated); widowed (including those persons living in common-law unions); separated (including those persons living in common-law unions); divorced (including those persons living in common-law unions); single (including those persons living in common-law unions). All the persons aged 15 years old and under are considered single.

2.7 Demographic adjustment to age structure

2.7.1 Demographic adjustment to age structure 0

The postcensal population estimate at age 0 almost entirely stems from the previous year's birth estimates. The estimate is therefore very reliable and can be used as a benchmark in adjusting the censal estimate. For 2011, the postcensal estimate at age 0 was first considered when modelling net census undercoverage. However, a demographic adjustment was required.

The demographic adjustment involves applying the sex ratio at age 0 of the postcensal estimates on the date of the census to the censal estimates for each province and territory. The adjustment concerns the number of men or women—whichever has the greatest error of closure—but could be made to both if the adjustment to be made to one is relatively greater.

The demographic adjustment reduces the error of closure at age 0 and makes it more equal by sex. The adjustment is then redistributed to other ages to maintain the total population by sex of the censal estimate for each province and territory.

2.7.2 Adjustments for older seniors

An analysis of the age structure of census counts and postcensal population estimates shows that populations of seniors over 85 years, and especially centenarians (population 100 years of age and older), were affected by an overestimation problem. For Canadian censuses from 1951 to 1991, an evaluation of the quality of reporting of older seniors reveals that overestimation problems affect men from 95 years of age and women from 100 years of age. The overestimation of centenarian populations was even more marked in the 1950s and 1960s (Bourbeau and Lebel 2000). In addition, the proportion of centenarians in Canada was shown to be greater than that of a set of countries with comparable death rates (Bourbeau and Lebel 2000; Kannisto 1999).

The problem of overestimating the number of centenarians, which seems to affect the Canadian census, are not unique to Canada. Specifically, such problems have been observed in the United States (Siegel and Passel 1976; Spencer 1987; Krach and Velkoff 1999; Kestenbaum and Ferguson 2005; Humes and Velkoff 2007; US Census Bureau 2012) and in the United Kingdom (Dini and Goldring 2008; Thatcher 1992, 1999). In the case of the very elderly, these problems can stem from errors related to voluntary reporting (from false declarations), involuntary reporting (from reporting errors, illiteracy or cognitive problems of the respondent, or from misreporting by proxy), data entry (manual or by scanner), questionnaire design, or from bias not detected by the data validation processes.

Census coverage studies that aim to adjust differential coverage, especially by age and sex, have trouble countering this phenomenon because of the small numbers of people surveyed beyond the age of 65, difficulty reaching seniors, and significant natural attrition (mortality) of these populations. The coverage studies of the last three censuses have made very few changes to the age structure of the base population in the population estimates beyond 85 years of age. To counter this phenomenon, manual validation procedures for the 2006 Census were implemented using a stratified sample of 95- to 99-year-olds and a complete selection of centenarians. An adjustment method for older seniors was also developed at the time of the previous rebasing in 2008, using the 2006 Census, and when the postcensal estimates from 2006 were produced.

However, when compared with the 2011 Census counts, the postcensal estimates of the number of centenarians based on the 2006 Census were 29% higher for women and 88% higher for men. This indicates that the downward adjustment made to the 2006 Census counts for the population aged 95 years and over—the population that would become the population of centenarians in 2011—was insufficient, or that the manual validation process for some questionnaires did not eliminate false nonagenarians and centenarians. This prompted Statistics Canada's Demographic Estimates Program to partially review its adjustment method for the age structure of the base population.

The new demographic adjustment for older seniors was made for the 2001, 2006 and 2011 censal estimates, and it reduced the number of people by 5% overall. The largest adjustments involved centenarians, reducing their counts on average by 28% for men and 4% for women. For the 95-to-99 age group, population counts were reduced by 13% for men and 3% for women. For example, with the new adjustments, the number of centenarians was revised downward in 2006 from 830 to 595 (-39%) for men and from 3,891 to 3,784 (-3%) for women. These revisions should not be interpreted as actual decreases in the number of older seniors, including centenarians, but rather as a new series of estimates based on a more robust method that offers more accurate estimates of populations of older seniors.

Using data on deaths between 1951 and 2011 and a combination of two methods (the method of extinct generations and the survival ratio method), the Demographic Estimates Program made a demographic adjustment to the age structure of the 2011 Census population, to prevent overestimation of the population of older seniors that was observed in the 2006 cycle. Assuming that populations of older seniors are hardly affected by migration, the basic principle of the method of extinct generations (Vincent 1951) is simple. Once all the individuals in a given generation have died, the historical number of survivors for this censal cohort, namely those born between May 10 in year x and May 9 in year x + 1, can be calculated for earlier years by inversely cumulating the number of deaths. The effectiveness of the method of extinct generations has been demonstrated many times by comparing populations reconstructed using this method with observed populations for countries with high-quality data (Coale and Kisker 1990; Coale and Caselli 1990; Kannisto et al. 1994; Human Mortality Database 2013). The file required to use this method comprises data on deaths by date of birth and date of death.

One drawback to this method is that the count for a given generation cannot be estimated exhaustively until all the individuals in that generation have died. An extinct generation is a cohort of individuals born in the same censal year who have all died by a given time (May 2011 in this case). This threshold is normally set at 110 years, meaning that generations for which death information is known through to the age of 110 will be considered extinct and complete. Since 1950, very few deaths have occurred beyond the age of 110 in Canada. This number has usually been 349 deaths in 10.6 million, or an average of one to two deaths per generation.

For non-extinct cohorts, the survival ratio method (Thatcher 1992; Thatcher et al. 2002; Andreev 2004) was used to produce an estimate based on the same principle, but according to the assumption that deaths in non-extinct cohorts are distributed by age like deaths in extinct cohorts. The population at a given age is estimated using the survival ratios of non-extinct cohorts, based on the ratios of the last five extinct cohorts. This method can be expressed as follows:

Equation 2.12: $$P_x^t = P_{x+1}^{t+1} + D_x^t$$

where P is the population and D is the deaths at age x at the start of census year t.

For a given year, the population (P) is calculated using the following equation:

Equation 2.13: $$P_x^T = \left(D_{x-1}^{T-1} + D_{x-2}^{T-2} + D_{x-3}^{T-3} + D_{x-4}^{T-4} + D_{x-5}^{T-5}\right) \times S_x^T \times c$$

where S is the survival ratio, T is the cohort and c is an adjustment factor that makes it possible for the ratio of older seniors to change over time (the modelled mortality is slightly reduced). This adjustment factor is based on the approach used by the United Kingdom's Office for National Statistics (ONS 2013). Adjustment factor c was set at 1.009 for men and 1.004 for women, and is based on the annual average change in life expectancy at 90 years of age for the last six extinct generations (that is, for people born between May 10, 1896 and May 9, 1901), by sex.

S is calculated as follows:

Equation 2.14:

$$S_x^T = \frac{\sum_{T-1}^{T-5} P_x}{\left(\sum_{T-2}^{T-6} D_{x-1} + \sum_{T-3}^{T-7} D_{x-2} + \sum_{T-4}^{T-8} D_{x-3} + \sum_{T-5}^{T-9} D_{x-4} + \sum_{T-6}^{T-10} D_{x-5}\right)}$$

The demographic adjustment was made by calculating the difference between the adjusted censal estimates and the estimate produced using the methods briefly discussed earlier for older seniors. To ensure the greatest consistency of estimates by cohort, the adjustment was made for the census populations of the last three censuses (2001, 2006 and 2011), by age and sex for each province and territory, starting at age 85 in 2001, age 90 in 2006, and age 95 in 2011. In addition, to minimize the impact on the age structure and on several indicators that base their calculations on population estimates (e.g., total fertility rate, life expectancy at birth, immigration rate by age group, etc.), surplus population counts, resulting from the difference between the censal estimates and the estimate of older seniors calculated using data on deaths, were redistributed among the population between 5 and 74 years of age based on the relative weight of the provinces and territories, by sex.

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 quality.[16]

Births are added to the base population, and deaths, excluding stillbirths, are subtracted (accounting as well for migration components) to generate the population estimates. The term births is commonly used to mean live births, while the term deaths refers to deaths excluding stillbirths.[17] This chapter contains information about the data sources for births and deaths and the methods of producing preliminary estimates.

3.1 Data sources and relevant concepts

Information on births and deaths is obtained from the vital statistics databases maintained by Statistics Canada's Health Statistics Division. The databases are created in collaboration with the provincial and territorial ministries responsible for recording vital statistics. Under provincial and territorial vital statistics acts (or equivalent legislation), the registration of all live births and all deaths is compulsory. The central vital statistics registry in each province and territory provides Statistics Canada with data on births and deaths. The data are available at the national, provincial, territorial, census division (CD) and census metropolitan area (CMA) levels.

The birth universe database includes births to Canadians, immigrants and non-permanent residents in Canada. The death universe database comprises deaths of Canadian residents, immigrants and non-permanent residents in Canada. Births and deaths of non-residents (visitors) are excluded.

3.2 Birth and death estimates, Canada, provinces and territories

3.2.1 Levels of estimates

There are three levels of estimates—preliminary, updated and final. These estimates are produced using vital statistics data provided by the Health Statistics Division. For final estimates, the numbers of births and deaths are taken directly from vital statistics provided by the Health Statistics Division (HSD). Vital statistics are received annually, with a lag of two to three years relative to our reference period. For preliminary and updated estimates, we use the ratio method described below.

3.2.2 Final estimates

Because of the completeness of the vital statistics databases, the data provided by the HSD for demographic estimates requires very few adjustments. When both place of birth or death and place of residence are available in the data files, place of residence is preferred for reporting the event for the purposes of population estimates as well as vital statistics health indicators. When a woman gives birth outside her province (i.e., in a province other than her province of residence), the birth is counted in the mother's province of residence. Similarly, if a person dies outside his or her province or territory of residence, the death is counted in the province or territory of residence. This is because the birth or death affects the population of the province of residence and not that of the province where the event occurred.

Using the component method, the number of births that occurred between the date on which the base population was established and the reference date is simply added to the base population and the number of deaths is subtracted from the base population.

16. Undercoverage of births due to late registration, births to Canadian residents outside Canada, non-registration of births, and deaths within days of birth is thought to be minimal. The same is true for undercoverage of deaths because of late registration, delayed or missing registrations due to unidentified bodies, and deaths of Canadians outside Canada. Deaths of Canadians serving in the Armed Forces outside Canada are not included in Statistics Canada's databases because they are not registered by the provinces and territories. Overcoverage of births and deaths is not considered significant and is not measured.
17. For a detailed description of the birth database and the death database maintained by the Health Statistics Division, refer to the following links on Statistics Canada's official website: Vital Statistics – Birth Database and Vital Statistics – Death Database.

3.2.3 Preliminary estimates[18]

Since vital statistics are not available at the time the preliminary population estimates are produced, births and deaths are estimated by applying probabilities to the population. The number of births is estimated using fertility rates by the mother's age group. The number of deaths is estimated using death rates by age group and sex. To produce preliminary or updated estimates, we use ratios based on the most recent year for which final estimates are available. This method is not used for Quebec or British Columbia since their statistical agencies always provide their most recent estimates (see below for details). The preliminary estimates are produced on a quarterly basis.

The following formulas are used to generate quarterly preliminary estimates of births and deaths at the provincial and territorial level.

Births:

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

Equation 3.1: $$B^{a_M}_{(t,t+2)} = \left(\frac{{}^{f}B^{a_M}}{{}^{f}P^{a_F}}\right) \times P^{a_F}_{t}$$

where

$(t,t+2)$ = interval of time representing the quarter that is being estimated;

$B^{a_M}_{(t,t+2)}$ = estimate of the number of births to be estimated between time t and t+2 to mothers M aged a;

${}^{f}B^{a_M}$ = number of final births f, for the corresponding quarter, from mothers M aged a;

${}^{f}P^{a_F}$ = estimate of the population of females F aged a at the beginning of the corresponding quarter for which final birth estimates f are available;

$P^{a_F}_{t}$ = estimate of the population of females F aged a at the beginning of the quarter for which an estimate is required (time t);

a = five-year age group, for women aged 15 to 49 years old.

Then the estimate of the total number of births is calculated as follows:

Equation 3.2: $$B_{(t,t+2)} = \sum_{a} B^{aM}_{(t,t+2)}$$

Deaths:

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

18. Unless otherwise noted, the term "preliminary" includes both preliminary and updated estimates when it refers to components.

Equation 3.3: $$D^{a,s}_{(t,t+2)} = \left(\frac{{}^{f}D^{a,s}}{{}^{f}P^{a,s}}\right) \times P^{a,s}_{t}$$

where

$(t,t+2)$ = interval of time representing the quarter that is being estimated;

$D^{a,s}_{(t,t+2)}$ = estimate of the number of deaths between time t and t+2 of people aged a and sex s;

${}^{f}D^{a,s}$ = number of final deaths f, for the corresponding quarter, people aged a and sex s;

${}^{f}P^{a,s}$ = estimate of the population of people aged a and sex s at the beginning of the corresponding quarter for which final death estimates f are available;

$P^{a,s}_{t}$ = estimate of the population of people aged a and sex s at the beginning of the quarter for which an estimate is required (time t);

a = five-year age group at the beginning of the quarter.

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

Equation 3.4: $$D_{(t,t+2)} = \sum_{s}\sum_{a} D^{a,s}_{(t,t+2)}$$

Monthly estimates are calculated using the most current final monthly distribution of births (or deaths).

The preliminary estimates cannot be finalized until the Health Statistics Division releases the vital statistics. Demography Division revises its estimates in the fall of each year.

3.2.4 Special treatment for preliminary estimates for Quebec and British Columbia

Quebec and British Columbia submit their most recent estimates of births and deaths to Statistics Canada's Demographic Estimates Program (DEP). Those estimates are based on counts available from their respective vital statistics offices. Their figures are used to produce preliminary or updated estimates for Statistics Canada's publication. For the final estimates, the two provinces' birth and death figures are taken from the vital statistics compiled by the Health Statistics Division.

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

Information about births and deaths from the Canadian vital statistics databases are used to produce other types of population estimates. For example, birth and death data (in addition to the other components of demographic growth) are used to produce the population estimates by age and sex, which will be discussed in Chapter 9.

Chapter 4

Immigration

This chapter provides information on the data sources pertaining to immigration, and the methods used to produce estimates of immigrants by age and sex 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 landed immigrants in Canada. An immigrant refers to a person who is or has ever been a landed immigrant (permanent resident) and who has been granted the right to live in Canada permanently by immigration authorities. Immigrants are either Canadian citizens by naturalization (the citizenship process) or permanent residents under Canadian legislation. Some immigrants have resided in Canada for a number of years, while others have arrived recently. Most immigrants are born outside Canada, but a small number are born in Canada. Also, children born in other countries to parents who are Canadian citizens that reside temporarily in another country are not included in the category as they become Canadian citizens at birth. For the Demography Division, the terms immigrant, landed immigrant and permanent resident refer to the same concept.

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 residents[19]: economic class, family class, and 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. CIC is currently transitioning towards a new system—the Global Case Management System (GCMS). The information is used to estimate the number and characteristics of people granted permanent resident status by the federal government for a given period. Immigrants are usually counted on or after the date on which they are granted permanent resident status or the right to live in 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 is straightforward, and adjustments to the data are not required. Information is available for each person entering Canada under immigrant status from CIC's administrative file.

Every month, CIC makes a data file containing the records of immigrants for the previous month available to Statistics Canada, 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 estimates, the file obtained from CIC indicates the province or territory of intended destination on 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 missing. These cases are distributed proportionally between the provinces and territories according to the observed distribution of immigrants for whom this information is available.

19. Children born abroad to Canadian parents who are out of the country are, by definition, Canadian citizens, and are, therefore, not included in immigration estimates. Included, however, are persons whose status changes from non-permanent resident (permit-holders, those with authorization, or refugee status claimants) to immigrant status from within Canada. Although their migration does not involve crossing the Canadian border, they are counted as non-permanent residents on their initial entry to Canada.

4.2.1 Immigration estimates by age and sex

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

4.2.2 Levels of estimates

The difference between preliminary and final estimates relies on how current the source used to estimate the component is. Since the FOSS file is continually being updated, new calculations are carried out every year to update the immigration estimates. Immigration estimates are preliminary for the first year, and finalized the following year.

Chapter 5

Net 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 who are living with them. This chapter provides details on the data sources for NPRs and the methods used to produce estimates of NPRs by age, sex, province and territory.

5.1 Data sources and relevant concepts

5.1.1 Data sources

NPR's population includes foreign workers, international students, holders of Minister's permits and refugees. This group also includes individuals who seek refugee status upon or after their arrival in Canada and remain in the country pending the outcome of processes relative to their claim. However, children born in Canada to parents with non-permanent resident status are considered Canadians by birth and have all the rights and privileges associated with citizenship. Note that Citizenship and Immigration Canada (CIC) uses the term temporary resident rather than non-permanent resident.

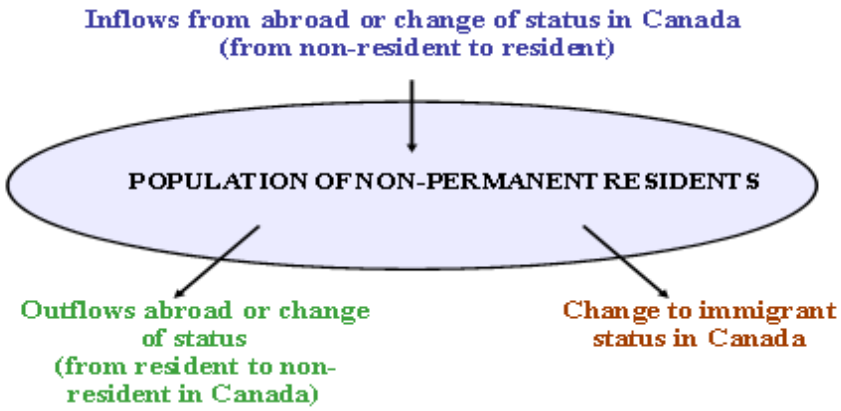
The data required to produce the NPR estimates are obtained from CIC Field Operations Support System (FOSS) files. CIC is currently transitioning towards a new system—the Global Case Management System (GCMS). FOSS files include data on visitor's permits, work permits, study permits, Minister's permits, refugee status claims, landings,[20] applications for landing and deportations. The information is used to estimate the number and characteristics of people granted non-permanent resident status by the federal government.

5.1.2 Relevant concepts

Movements into and out of the NPR population are referred to as flows, as are status changes from permit holders to refugee status claimants or from permit holders or refugee status claimants to immigrant status.[21] As illustrated in Figure 5.1, persons can enter into the current stock of NPRs (inflows) either from abroad or from non-resident status within Canada (e.g. a visitor). When people leave the NPR population (outflows), it is assumed that they have either left the country, become a non-resident, or become an immigrant.

20. Landings refers to lawful permission to establish permanent residence in Canada.
21. Persons leaving the NPR population by obtaining immigrant status change the size of the NPR population, but this does not have an impact on the size of the total population of Canada since these persons become a part of the permanent resident population.

Figure 5.1
Non-permanent residents flows



In the Demographic Estimates Program, the number of NPRs is included in the total population estimates for a given date, while the difference in NPR numbers over a certain period of time is included as a component of demographic growth. This new component is referred to as net NPRs.

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

i. as holders of permits (they may concurrently hold more than one type);

ii. as refugee status claimants;

iii. as refugee status claimants who also possess one, two or three of the types of permits that would qualify them for NPR status.

In cases where refugee status claimants concurrently hold one of the other permits (for instance if a refugee status claimant is granted an employment permit to help support himself or herself along with dependants within Canada), the refugee status supersedes the other NPR status. Consequently, references to permit holders excludes persons who have also made refugee status claims. The two major NPR sub-groups—permit 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 included, this population being covered by the birth component (see Chapter 3).

All persons in possession of a permit or claiming refugee status are assigned a Client Identification Number (CID) by CIC. This identification number is a key variable as it allows identifying only once each person holding a permit or claiming refugee status regardless of the number or permits issued to the holder. The CID 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 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 the amount of non-permanent residents

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

5.2.1 Number of permit holders

Permit 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. A person is thus considered part of this population if he or she possesses a valid work permit, study permit or Minister's permit on the estimate reference date. 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, the information from the permit for which the effective date[22] is closest to the reference date is used.

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

Status as a permit holder terminates upon the expiry of the valid document or when a PH is granted immigrant status (i.e., becoming a permanent resident), deported or seeks refugee status.[23] 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, then the province of residence is taken from the permit for which the effective date is closest to the reference date. In some circumstances, the province of residence is not indicated on the document. For cases where this information is missing, CIC has developed an alternative way to attribute a province or territory of residence by using other sources of information. All remaining 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 and territory, the number 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 Amount of refugee status claimants

A person is considered part of the refugee status claimant (RSC) population once a claim has been filed at a Canada 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 (a separate case file is opened for each claimant, including dependants, and is assigned a CID).

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 landed 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.

For RSCs seeking immigrant status:

i. it is assumed that they will leave the NPR population 2 years after the date of their application for landing 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;

22. The effective date is the date (exact or approximate) when the PH entered Canada as an NPR, or the date as of which a permit is extended

23. At times, there may appear to be a short interruption in a PH's temporary stay in Canada. It could be due to fact that a permit expires and there is a brief time lag between the expiry date and the effective date of the next permit. Since this is likely due to administrative delays in the issuance of permits and extensions, interruptions of less than 31 days are disregarded and these persons are considered as having been continuously residing in Canada.

ii. it is assumed that if they do not have valid proof of identity[24] 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 CIC.

For RSCs not seeking 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 to request a status update, file a new application, 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 the RSC is no longer an NPR because he or she has either left the country or is deceased.

The province of residence for each RSC may be obtained from the claim record. If the information is missing or invalid, when possible it is extracted from the permit entry date that is closest to the reference date. The province of residence might not always be indicated in the document. For cases where this information is missing, CIC has developed an alternative way to attribute a province or territory of residence by using other sources of information. All remaining cases are prorated according to the provincial distribution for RSCs where the province of residence is available (or derived).

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

For each province and 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[25] and final estimates depends on the timeliness of the source used to estimate this component. Estimates of the number and net change of NPRs are revised on an annual basis. Non-permanent resident estimates are preliminary the first year and updated the following year. They become final two to three years after the reference year.

5.3 NPR estimates by age and sex

5.3.1 Permit holders

Though permits include information on the number of persons the document covers, data related to the age and sex 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 (one application per person), the age and sex 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 and age for these claimants are obtained directly from their valid claim.

24. 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 proof of 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.
25. Unless otherwise noted, the term "preliminary" includes both preliminary and updated estimates when it refers to components.

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

The net change of NPRs for a given period is calculated by subtracting the amount of NPRs at the beginning of the period from the amount at the end of the period.

Equation 5.1: $$\Delta NPR_{(t,t+i)} = NPR_{(t+i)} - NPR_t$$

where

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

$\Delta NPR_{(t,t+i)}$ = change in the number of NPR between time t and time t+i;

$NPR_{(t+i)}$ = number of NPR at time t+i;

NPR_t = number of NPR at time t.

In the Demographic Estimates Program, net change in NPRs is one of the components added to the base population to produce population estimates on a given date.

Chapter 6

Emigrants, net temporary emigrants and returning emigrants

Emigration refers to citizens or immigrants who leave the country to establish a residence in another country. This residence abroad may be permanent or temporary. Emigration results in a loss of population. Taking into account all aspects of Canadian emigration, the calculation of population involves the estimation of emigrants, net temporary emigrants 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 the United States.

As emigration components are the most difficult components 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 nation's 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 persons establishing a permanent residence in another country (i.e., emigrants), persons living temporarily abroad (i.e., net temporary emigrants), and finally the portion of emigrants who have returned to Canada (i.e., returning emigrants). Different data sources and methods are necessary for the two types of emigrants, as well as for returning emigrants. 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 immigrants who have left Canada to establish a permanent residence in another country (sometimes referred to as permanent emigration). For example, persons 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. 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 Office of Immigration Statistics of U.S. Department of Homeland Security provides data on Canadians who acquire permanent 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 (CCTB) program and tax data from Canada Revenue Agency (CRA) is used.[26] For this portion of emigration, we cannot determine if someone has acquired permanent immigrant status in another country. In these cases, we must therefore assume that someone who has cut ties with Canada, according to CRA data, has established permanent residence in another country.

Some people leave Canada to live temporarily in another country while not maintaining a usual place of residence in Canada. Other temporarily leave Canada and then 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 National Household Survey (NHS), combined with Demography Division's estimates of returning emigrant, are used to estimate the number of temporary emigrants returning.

26. The T1 Family File (T1FF) is derived from the Canada Revenue Agency (CRA) T1 file by Income Statistics Division of Statistics Canada.

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, a 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 as temporary residents abroad and their consecutive returns to Canada has an important effect on the quality of the country's population estimates (Michalowski 1999). Including net change in temporary emigration in the Demographic Estimates Program is a practice that started in 1998, with revisions to estimates back to 1991.

Returning emigrants are Canadian citizens or immigrants having previously emigrated from Canada and subsequently returned to Canada to re-establish a permanent residence. Again, data from the CCTB program and from CRA's T1FF are used in estimating returning emigrants.

6.2 Estimates of emigrants, net temporary emigrants and returning emigrants

6.2.1 Emigrants

The number of emigrants is estimated using data from the Office of Immigration Statistics, U.S. Department of Homeland Security, data collected by the Canada Child Tax Benefit (CCTB) program, and data from the T1 Family File (T1FF). The first source is used to estimate emigration to the United States while CCTB data is used to estimate emigration to other countries. The CCTB data covers all emigrant children, thus we can use this data to derive the children and adults emigrating to a country other than the United States. The estimates of the number of child emigrants have to be adjusted because the CCTB program is not universal. As a result, four adjustment factors are used to take into account:

1. the incomplete coverage due to a delay in the receipt and processing of the files of children eligible for the CCTB. Emigration estimates are usually finalized two or three years after the reference period. Based on historical files, it seems that CCTB data recordings reached a significant level two years after the year of emigration and that it takes four years for CCTB files to become complete. The data is adjusted as it is used within this four-year time frame;
2. the program's partial coverage, that is, people who do not apply for the CCTB or are not eligible;[27]
3. the differential propensity to emigrate between children as a whole and those who are eligible for the CCTB; and
4. the differential propensity to emigrate between adults and children.

To calculate these adjustment factors, we make use of CCTB and tax data from T1FF. The estimation methods for child and adult emigrants are described in the following section.

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

Emigrant children

The CCTB file provides the numbers of dependent children (under 18 years of age) whose parents are eligible and have applied for the CCTB, but are no longer residents as defined by the CRA. These data are available on a monthly basis, but only provided by CRA to Statistics Canada annually for each province and territory. As mentioned, the CCTB program is not universal; therefore the data is incomplete in terms of estimation of total child emigrants and requires adjustments.

The first correction factor, that is the adjustment for partial coverage, corrects for the absence of universality of the CCTB program. 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 correction factor uses 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, province and territory.

The second correction factor is used to adjust the differential emigration propensity between all children, and children registered for the CCTB program for each province and territory, on an annual basis. This factor is obtained by comparing the emigration rates for all children (aged 0 to 17) with the rates of CCTB-eligible children. This factor is calculated for each province and territory and is based on the last three available years of the T1FF. To eliminate variations due to small numbers in each of the Atlantic provinces, 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 third correction factor is to take into account the incompleteness of the emigration information from the CRA's CCTB data files due to delays in the recordings of an emigration. Based on comparisons done with files that were two, three and four years after the reference period, it appeared that the CRA files could be considered complete about four years after the reference period. As we use CCTB data within this four year timeframe, we apply a correction factor. The same factor is applied to the monthly data of each province or territory.

The formula to estimate child emigrants is as follows:

For each province and territory:

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

where

${}_{j}EM^{0-17}$ = adjusted emigrant children aged 0 to 17 from province or territory j;

${}^{CCTB}_{j}EM^{0-17}$ = number of emigrant children according to CCTB moving from province or territory j;

${}^{CCTB}_{j}\mathfrak{R}$ = coverage rate of CCTB program for province or territory j;

${}_{j}G$ = adjustment factor for emigration propensity of non CCTB children (derived from tax data), for all provinces, except for the Atlantic provinces, where the factor is measured as a whole and in the territories where the Canadian factor is used;

D = adjustment factor for the incompleteness of emigration data from CCTB files due to delays in the recordings of an emigration.

The factors used in the equation above are detailed in equations 6.2 to 6.4c: factors for incomplete coverage of CCTB-registered children (equation 6.2), factors for the differences in the propensities of CCTB-registered and non-registered subpopulations to emigrate (equation 6.3), as well as factors used to correct the incompleteness of emigration data due to delays in the recordings of an emigration (equations 6.4a to 6.4c).

The coverage rates are calculated on a monthly basis as follows:

For each province and territory:

Equation 6.2: $${}^{CCTB}_{j}\mathfrak{R} = \left[\frac{{}^{CCTB}_{j}P^{0-17}}{{}^{Dem}_{j}P^{0-17}}\right]$$

where

${}^{CCTB}_{j}\mathfrak{R}$ = coverage rate of CCTB program for each province or territory j;

${}^{CCTB}_{j}P^{0-17}$ = number of children aged 0 to 17 years registered for the CCTB program in each province or territory j;

${}^{Dem}_{j}P^{0-17}$ = population estimated by Demography Division of children aged 0 to 17 years in each province or 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:[28]

Equation 6.3: $${}_{j}G = \frac{{}^{TAX}_{j}RateEM^{0-17}}{{}^{CCTB\_TAX}_{j}RateEM^{0-17}} = \frac{\left[\frac{{}^{TAX}_{j}EM^{0-17}}{{}^{TAX}_{j}P^{0-17}}\right]}{\left[\frac{{}^{CCTB\_TAX}_{j}EM^{0-17}}{{}^{CCTB\_TAX}_{j}P^{0-17}}\right]}$$

where

${}_{j}G$ = adjustment factor for emigration propensity of non-CCTB children (derived from fiscal data) for each province of origin j;

${}^{TAX}_{j}RateEM^{0-17}$ = emigration rate of all children from each province j;

${}^{CCTB\_TAX}_{j}RateEM^{0-17}$ = emigration rate of CCTB-registered children from each province j, derived from T1FF;

${}^{TAX}_{j}EM^{0-17}$ = emigration of all children from province j, according to income tax data;

${}^{TAX}_{j}P^{0-17}$ = ISD estimate of children in province j, based on income tax files;

${}^{CCTB\_TAX}_{j}EM^{0-17}$ = emigration of CCTB-registered children from province j, derived from the T1FF by ISD;

${}^{CCTB\_TAX}_{j}P^{0-17}$ = CCTB-registered children in the population of the income tax file provided by ISD by province j.

We use delay factors to compensate for the incompleteness of the CCTB data file on child emigration due to registration delays.

Table 6.1 illustrates the calculation of the delay factors; all numbers in the table are fictitious. In this table, we suppose that the production year is 2011, therefore the t-3, t-2 emigration year in the first column corresponds to 2008/2009 and the t-4, t-3 emigration year corresponds to 2007/2008, etc. The subsequent three columns of the table show numbers of child emigrants by year of emigration and by duration between the emigration year and the year of the data file used. In the table, the numbers of child emigrants for the emigration year t-4, t-3 from the data

28. Atlantic provinces all have the same factor or rate and the Canadian factor or rate is used for the territories.

file that is two years after the emigration year is 7,900 (duration of two years between the emigration year and the data file year) and the number is 8,500 from the data file that is three years after the emigration year (duration of three years between the emigration year and the data file year). Our objective is to estimate what the t-4, t-3 and also the t-3, t-2 emigration year numbers would be in a file that is four years after the emigration year.

Table 6.1
Calculation of the delay factors: An illustration for production year 2011

	Number of child emigrants in the CCTB file after					
Year of emigration	2 years	3 years	4 years	$D_{(2,3)}$	$D_{(3,4)}$	$D_{(2,4)}$
t-7, t-6 (2004-2005)	6,300	6,750	7,100	1.071	1.052	1.138[2]
t-6, t-5 (2005-2006)	6,700[1]	7,250	7,600	1.082	1.048	1.142
t-5, t-4 (2006-2007)	7,200	7,850	8,200	1.090	1.045	1.124
t-4, t-3 (2007-2008)	7,900	8,500	Z=8,908	1.076		
t-3, t-2 (2008-2009)	7,700		Y=8,740			
Average factor (3 years)				1.083	1.048	1.135

1. The data in diagonal are from the same data file year, as indicated by the different shades of grey. Darkest grey represents data from the 2008 file while lightest grey relates to 2011 data.
2. To obtain $D_{(2,4)}$, one multiplies $D_{(3,4)}$, which is on the same line, by $D_{(2,3)}$, which is on the next line. For example, the $D_{(2,4)}$ value: 1.138 is equal to $D_{(3,4)}$: 1.052 times $D_{(2,3)}$: 1.082.

We first calculate delay factors for the previous emigration years. The delay factors (D) in the right-hand columns are derived from the numbers on the left. $D_{(2,3)}$ is the ratio between duration 3 and duration 2; while $D_{(3,4)}$ is the ratio between duration 4 and duration 3. For example, $D_{(3,4)}$ of 1.045 for the emigration year t-5, t-4 comes from the ratio between 8,200 and 7,850. This factor could be used to project one year ahead from duration 3 to duration 4. But, to project two years ahead from duration 2 to duration 4, we calculate the factor D(2,4) (the last column in Table 6.1) by multiplying the two most recent factors, $D_{(2,3)}$ and $D_{(3,4)}$.

The formulas are as follows:

Equation 6.4a:
$$D_{(2,3)} = \frac{{}^{CCTB}EM^{t}_{(t-4,t-3)}}{{}^{CCTB}EM^{(t-1)}_{(t-4,t-3)}}$$

Equation 6.4b:
$$D_{(3,4)} = \frac{{}^{CCTB}EM^{t}_{(t-5,t-4)}}{{}^{CCTB}EM^{(t-1)}_{(t-5,t-4)}}$$

where

$${}^{CCTB}EM^{t}_{(t-x-1,t-x)}$$ = the number of child emigrants for the emigration year (t-x-1, t-x) from the CCTB file of year t.

If we want to project numbers for an emigration year that is three years before the production year, then we use the delay factor in the recordings of data on emigration from duration 3 to duration 4 (i.e., $D_{(3,4)}$).

If we want to project numbers for an emigration year that is two years before production year, then we use:

Equation 6.4c:
$$D = D_{(2,4)} = D_{(2,3)} \times D_{(3,4)}$$

In production, we use a three-year average of each of the $D_{(3,4)}$ and the $D_{(2,4)}$ factors based on the most recent factors available. For example, starting from data in table 6.1, in year t, we will project the emigration number for the year t-3, t-2 (7,700) with a three-year average of the most recent $D_{(2,4)}$. The average is based on the factors: 1.138, 1.142 and 1.124. This average factor (1.135) will then be used to multiply the number of emigrants in year t-3, t-2 at duration 2 (7,700) and the results (Y=8,740) will be the number of expected CCTB child emigrants in the file that is four years after the emigration year. The same average calculation of the $D_{(3,4)}$ factor (1.048) is used to multiply the

emigration number of year t-4, t-3 at duration 3 (8,500) to obtain the expected number of CCTB child emigrants (Z=8,908) in the file that is four years after the emigration year.

Data on Canadian adults and children who became immigrants in the United States are available from the Department of Homeland Security on an annual basis. The Homeland Security data provides quarterly flow of immigrants to the United States emigrating from Canada. We then subtract this number of child emigrants to the United States from the estimated total number of child emigrants from CCTB files to produce an estimated number of child emigrants to countries other than the United States. The calculations can be derived as follows:

Equation 6.5: $${}^{\mathrm{OC}}\mathrm{EM}^{0-17} = {}^{\mathrm{CCTB}}\mathrm{EM}^{0-17} - {}^{\mathrm{HS}}\mathrm{EM}^{0-17}$$

where

${}^{\mathrm{OC}}\mathrm{EM}^{0-17}$ = number of emigrant children to countries other than the United States;

${}^{\mathrm{CCTB}}\mathrm{EM}^{0-17}$ = number of emigrant children based on CCTB data;

${}^{\mathrm{HS}}\mathrm{EM}^{0-17}$ = number of emigrant children to the United States based on United States Department of 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 emigrant adults 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 countries.

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

Equation 6.6: $${}^{\mathrm{OC}}\mathrm{EM}^{18+} = \frac{{}^{\mathrm{OC}}\mathrm{EM}^{0-17}}{{}^{\mathrm{Dem}}\mathrm{P}^{0-17}} \times {}^{\mathrm{Dem}}\mathrm{P}^{18+} \times \mathrm{RatioAC}$$

where

${}^{\mathrm{OC}}\mathrm{EM}^{18+}$ = number of monthly adult emigrants to countries other than the United States;

$\frac{{}^{\mathrm{OC}}\mathrm{EM}^{0-17}}{{}^{\mathrm{Dem}}\mathrm{P}^{0-17}}$ = monthly emigration rate of children moving to countries other than the United States;

${}^{\mathrm{Dem}}\mathrm{P}^{18+}$ = population estimate of adults at the beginning of the month;

$\mathrm{RatioAC}$ = annual ratio of adult emigration rate and child emigration rate to countries other than the United States.

The $\mathrm{RatioAC}$ is the rate of adult emigration divided by the rate of child emigration to countries other than the United States taken from tax data.[29] It is calculated as follows:

29. The tax data identifies the destination of emigration for the taxfilers who provided their address in another country. Emigrants who provided a Canadian address on their tax form are proportionally distributed between the United States and countries other than the United States. Emigration is defined by the departure date on the tax form which indicates the day when a person is no longer a resident of Canada according to the CRA.

Equation 6.7: $$\text{RatioAC} = \frac{\left[\frac{\text{TAX}_{\text{EM}^{18+}}}{\text{Dem}_{\text{P}^{18+}}}\right]}{\left[\frac{\text{TAX}_{\text{EM}^{0-17}}}{\text{Dem}_{\text{P}^{0-17}}}\right]}$$

where

$\text{TAX}_{\text{EM}^{18+}}$ = annual number of adult emigrants to countries other than the United States, measured from tax data ;

$\text{Dem}_{\text{P}^{18+}}$ = estimate of adult population by Demography Division (based on the average of start-of-period and end-of-period populations);

$\text{TAX}_{\text{EM}^{0-17}}$ = annual number of child emigrants to countries other than the United States, measured from tax data;

$\text{Dem}_{\text{P}^{0-17}}$ = estimate of child population by Demography Division (based on the average of start-of-period and end-of-period populations).

Finally, the total number of adult emigrants corresponds to the following:

Equation 6.8: $$\text{EM}^{18+} = {}^{\text{OC}}\text{EM}^{18+} + {}^{\text{HS}}\text{EM}^{18+}$$

where

EM^{18+} = adult emigrants;

${}^{\text{OC}}\text{EM}^{18+}$ = number of adult emigrants to countries other than the United States;

${}^{\text{HS}}\text{EM}^{18+}$ = number of adult emigrants to the United States, according to U.S. Department of Homeland Security data.

The data provided by the U.S. Department of Homeland Security does not include the province of origin for Canadian emigrants. The estimates for adults' province of origin at the provincial and territorial level must be acquired from another source. The ratio between the number of adult and child emigrants provided by tax data (T1FF), is used to estimate the distribution of adult emigrants by province and territory of origin.

The provincial and territorial of adult migrants is estimated as follows:

Definitions of the symbols used:

$\text{TAX}_{\text{j}}\text{EM}^{18+}$ = emigrants from province or territory j, aged 18 years or more, taken from tax data (T1FF);

$\text{TAX}_{\text{j}}\text{EM}^{0-17}$ = emigrants from province or territory j, aged 0 to 17 years, taken from tax data (T1FF);

$\text{CCTB}_{\text{j}}\text{EM}^{0-17}$ = emigrants of province or territory j, aged 0 to 17 years, estimated from CCTB;

${}^{\text{W}}_{\text{j}}\text{EM}^{18+}$ = estimate of emigrants from province or territory j, aged 18 years or more which will serve to calculate the provincial and territorial distribution of adult emigrants.

The method is composed of three steps:

a. To estimate the number of adult emigrants by province and territory, the ratio of adult to child emigrants from the T1FF is applied to the estimate of children from the CCTB

Equation 6.9:

$$\text{If } {}^{\text{TAX}}_{j}\text{EM}^{0-17} \text{ or } {}^{\text{CCTB}}_{j}\text{EM}^{0-17} = 0 \quad \text{then} \quad {}^{\text{W}}_{j}\text{EM}^{18+} = {}^{\text{TAX}}_{j}\text{EM}^{18+}$$

If not

$${}^{\text{W}}_{j}\text{EM}^{18+} = \left[\frac{{}^{\text{TAX}}_{j}\text{EM}^{18+}}{{}^{\text{TAX}}_{j}\text{EM}^{0-17}}\right] \times {}^{\text{CCTB}}_{j}\text{EM}^{0-17}$$

Due to small results, for each of the individual territories the fraction in equation 6.9 is replaced by the national fraction. For Prince Edward Island, the fraction is replaced by Atlantic Provinces fraction.

b. The distribution by province and territory is then calculated for the estimated number of adult emigrants

Equation 6.10: $$\text{W}_j = \frac{{}^{\text{W}}_{j}\text{EM}^{18+}}{\sum_j {}^{\text{W}}_{j}\text{EM}^{18+}}$$

W_j represents the weight of each province and territory and establishes the provincial and territorial distribution of adult emigrants

c. The distribution is then applied to the total adult emigrants (EM^{18+} from equation 6.8)

Estimates of emigrants by age and sex

The distribution by age, sex, province and territory of emigrants is taken from the emigrants in the tax data (T1FF) adjusted for the variability of the T1FF coverage by age, sex, province and territory. The adjustment is as follows:

Equation 6.11:

$$\text{If } {}^{\text{T1FF}}_{a,s,j}\text{P} = 0 \quad \text{then} \quad {}^{\text{T1FF}-\text{adj}}_{a,s,j}\text{EM} = 0$$

If not

$${}^{\text{T1FF}-\text{adj}}_{a,s,j}\text{EM} = {}^{\text{T1FF}}_{a,s,j}\text{EM} \times \left(\frac{{}^{\text{EST}}_{a,s,j}\text{P}}{{}^{\text{T1FF}}_{a,s,j}\text{P}}\right)$$

where

${}^{\text{T1FF}-\text{adj}}_{a,s,j}\text{EM}$ = number of T1FF emigrants adjusted for coverage based on age, sex and province or territory;

${}^{\text{T1FF}}_{a,s,j}\text{EM}$ = number of T1FF emigrants based on age, sex and province or territory;

${}^{\text{EST}}_{a,s,j}\text{P}$ = Demography Division's population estimates by age, sex and province or territory on January 1st of the year following the T1FF year;

${}^{\text{T1FF}}_{a,s,j}\text{P}$ = T1FF population by age, sex and province or territory.

The adjusted number of emigrants from the two most recent years of T1FF data are assembled into five year age groups for each province, territory and sex. Due to small numbers, Atlantic province numbers are grouped together and used for the each Atlantic province. The National numbers are used for each of the territories.

Using the Sprague coefficients, we then produce the numbers by age, sex, province and territory.[30] We bring back the age of a person, which is on December 31 on the T1FF to their age at July 1. These numbers represent the emigration distributions for the year that covers the beginning of July (x) to the end of June (x+1), by age and sex for each province and territory. Therefore if the distribution estimate was produced by combining the 2010 and 2011 T1FF, we apply the distributions to the year beginning in July 2010 and ending in June 2011.

Levels of estimates

The difference between preliminary[31] and final estimates lies in the timeliness of the sources used to estimate this component. The same estimation method is used.

6.2.2 Net temporary emigration

Some people leave Canada to temporarily live in another country; others who were temporarily outside Canada return. The net result of those departures and returns is the component known as net temporary emigration. The following steps are used to estimate their monthly numbers by province or territory:

1. The 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 temporarily[32] during an intercensal period and who are still out of the country at the end of the period;
2. The 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 National Household Survey (NHS). The NHS 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 (NHS estimates Step 2a) we subtract Demography Division's estimate of returning emigrants. The resultant estimate is the number of temporary emigrants returning;
3. The estimate of net temporary emigration for Canada is derived by subtracting returning persons (Step 2b) from departures (Step 1);
4. The estimate derived in Step 3 is then distributed by province or territory according to the provincial and territorial distribution of departures of temporary emigrants based on the RRC.[33] The number for the Atlantic provinces is estimated as a group and redistributed proportionately to each province according to their respective population size. The same is done for the three territories;
5. 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 6.12: $$\mathrm{NTE}_m = \frac{\left[\left(\frac{\mathrm{NTE}}{12}\right) + \left(\frac{\mathrm{EM}_m}{\mathrm{EM}} \times \mathrm{NTE}\right)\right]}{2}$$

where

NTE_m = number of net temporary emigrants for the month m;

NTE = annual number of net temporary emigrants;

30. Sprague's multipliers are interpolation coefficients used to subdivide data. For a detailed description of the method of Sprague's multipliers, see Shryock, Siegel et al. 1976.
31. Unless otherwise noted, the term "preliminary" includes both preliminary and updated estimates when it refers to components.
32. Emigration is defined as temporary based on the intention of returning and the time spent outside the country
33. Special estimates are done for the territories as the RRC does not provide estimates.

EM_m = number of emigrants for the month m;

EM = number of annual 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 and sex

To obtain the estimates of net temporary emigrants by age and sex, the same distributions as observed for emigrant are used.[34]

Levels of estimates

The difference between preliminary and final estimates lies in the timeliness of the emigration estimate used to calculate the seasonal adjustment for the net temporary emigration. The same estimation method is used.

6.2.3 Returning emigrants

A returnig emigrant is a person returning to Canada, after having been classified as an emigrant. In a manner similar to the procedure used to calculate the number of emigrants, data from the CCTB file and CRA's T1FF are used to estimate the number of returning emigrants. A citizen or an immigrant 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.

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 for the reason of incompleteness of data on births. As is the case with emigrants, 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 numbers of dependent children (under the age of 18 years) of CCTB recipients who have returned to Canada after a period of emigration. These data are available on a monthly basis but provided by CRA to Statistics Canada annually for each province and territory.

As with emigrant children, adjustment factors are applied to adjust the CCTB data. First, the CCTB data for returning child emigrants are adjusted with a factor reflecting the program's partial coverage, that is, people who do not apply for the CCTB or are not eligible. This factor is obtained by comparing the estimated number of children in the population with the number of children in CCTB files (equation 6.2). The second adjustment factor is used to take into account the differential propensity to emigrate between children who are eligible for the CCTB and children who are not. In this case, it is assumed that the ratio is the same as the ratio for emigrants (equation 6.3). Hence, the numbers for returning CCTB-eligible children are adjusted with the same differential propensity factors that are used for emigration. Finally, we adjust the data for the delay in completion of CCTB data (the delay factor), which is calculated the same way as was done for emigration, but this time with the number of returning emigrants. The formula provided earlier used to compute the estimate of child emigrants (equation 6.1) is used to estimate returning child emigrants, where $\left({}^{CCTB}_{j}EM^{0-17}\right)$, the number of child emigrants by province or territory, according to CCTB data, is replaced by $\left({}^{CCTB}_{j}RE^{0-17}\right)$, the number of children returning to Canada.

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

Returning emigrant adults

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

Equation 6.13: $$\tau^{RE} = \frac{{}^{C}RE^{18+}}{{}^{C}RE^{0-17}}$$

where

τ^{RE} = ratio of adults and children in the returning emigrant population;

${}^{C}RE^{18+}$ = number of returning emigrant adults according to NHS data;

${}^{C}RE^{0-17}$ = number of returning emigrant children according to NHS data.

The number of adult returning emigrants is estimated by multiplying the estimate of returning child emigrants based on CCTB data by the adult to child ratio from equation 6.13.

Equation 6.14: $$RE^{18+} = {}^{CCTB}RE^{0-17} \times \tau^{RE}$$

where

RE^{18+} = estimated number of returning adult emigrants;

${}^{CCTB}RE^{0-17}$ = estimated number of returning child emigrants according to CCTB data;

τ^{RE} = Adult-child ratio for the returning emigrant population according to NHS counts.

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 and sex

The age and sex distribution of returning emigrants is based on NHS data. Characteristics of returning emigrants are derived from census mobility data for one year, after excluding non-permanent residents and immigrants. Since 2011-2012, we use a distribution by age and sex derived from the 2011 NHS. The NHS distribution by single years of age and sex is applied to children aged 0 to 17 years and to adults 18 years and over.

Levels of estimates

The difference between preliminary and final estimates lies in the timeliness of the sources used to estimate this component. The same estimation method is used.

Chapter 7

Interprovincial migration

Interprovincial migration represents movement between provinces or territories involving a change in the usual place residence. Intraprovincial migration (or subprovincial migration) also involves a change in usual place of residence, but these movements occur within the same province or territory. The term internal migration may refer to either interprovincial migration or intraprovincial migration, or both. This chapter focuses on the interprovincial migration component. For information on intraprovincial migration, see Chapter 8 on subprovincial population estimates.

As is the case for emigration, there is no provision for recording interprovincial migration in Canada. Consequently, this component of demographic growth has to be estimated using administrative data. The methods used to prepare estimates of interprovincial migration will be discussed in this chapter.[35]

7.1 Relevant concepts and data sources

Interprovincial migration is the movement of persons from one province or territory to another, involving a change to their usual place of residence. These movements are measured by comparing the place of residence at the beginning and end of the measurement period.

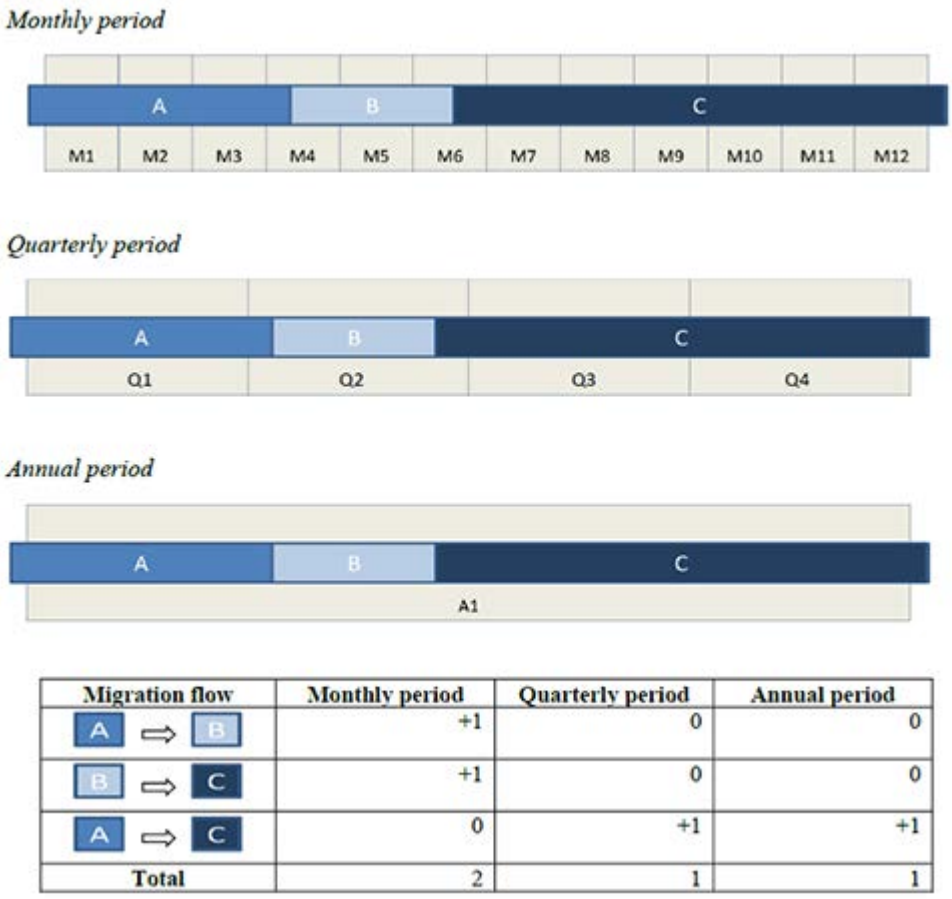
Canada Revenue Agency (CRA) provides Statistics Canada with data from the Canada Child Tax Benefit (CCTB) program and personal income tax returns, which are used to estimate interprovincial migration. Monthly, quarterly and annual interprovincial migration estimates are produced using these sources.

Since 1976, personal income tax records have become the official data source for final interprovincial 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. Statistics Canada receives tax data only after annual processing, compared to CCTB data which are available on a monthly basis. Therefore, final estimates are calculated annually using the T1 Family File (T1FF),[36] while preliminary estimates[37] are calculated using CCTB data.

Interprovincial migration is a unique demographic phenomenon, in that it is closely tied to the measurement period. In fact, as a result of return migration and multiple migrations, the number of migrants may differ substantially depending on the duration of the measurement period. The following diagram illustrates this situation using a case of multiple migration.

35. Starting with the release dated July 1, 2015, interprovincial migration estimates will be produced using a new method. This method was used to revise interprovincial migration estimates from July 2011 onward. Details on this new method can be found in this document. For information on the previous method, please refer to the 2012 edition of this publication.
36. The T1 Family File (T1FF) is derived from the CRA's T1 file by Statistics Canada's Income Statistics Division. In this document, "T1FF" is used to refer to this file.
37. Unless otherwise noted, the term "preliminary" includes both preliminary and updated estimates when it refers to components.

Figure 7.1
Mock migration flow of a migrant starting the year in province A and ending it in province C, by measurement period



Migration flow	Monthly period	Quarterly period	Annual period
A ⇨ B	+1	0	0
B ⇨ C	+1	0	0
A ⇨ C	0	+1	+1
Total	2	1	1

Even though the migrant moves twice in the same year, only one move is recorded at the annual level (A to C). As well, since both moves occur in the second quarter, only one move is recorded at the quarterly level (A to C). Lastly, since the moves occur in the fourth and sixth months, two moves are recorded at the monthly level (A to B and B to C). The result is different migration flows depending on the measurement period.

Since interprovincial migration estimates are calculated for the month, quarter and year, the number of migrants must be calculated for each measurement period. Although the number of migrants may differ, net interprovincial migration[38] must be identical for each measurement period. For example, annual net migration in the figure above is -1 for A, 0 for B and +1 for C regardless of whether it is calculated using monthly periods or quarterly periods. However, given the nature of the administrative data and the fact that some individuals are absent at the beginning and end of each period, an adjustment must be made to ensure that the net migration is identical.

7.2 Preliminary interprovincial migration estimates, Canada, provinces and territories

7.2.1 Estimation methodology

Preliminary estimates of migration between provinces and territories are produced using CCTB data, along with data from personal income tax records, processed by Income Statistics Division. Because the CCTB program is

38. Net interprovincial migration is the difference between in-migrants and out-migrants for a given province or territory.

not universal and does not provide direct information on the number of adult migrants, this estimation requires three adjustment factors to derive the complete estimate of interprovincial migration:

1. the program's partial coverage (e.g., those who have not applied for the CCTB);
2. the differential propensity to migrate between children who are registered to the CCTB program and all children;
3. the differential propensity to migrate of adults versus children.

The three adjustment factors use CCTB and T1FF data. The methods for estimating the interprovincial migration rates of children (aged 0 to 17 years) and adults (aged 18 years and older) are described in the next section.

Interprovincial migration of children

Given that income tax returns are produced annually and are not available when preliminary estimates are being produced, preliminary estimates of interprovincial migration are based on monthly microdata from CCTB administrative files, from which the number of child migrants can be extracted. The data can be used to derive information on the usual province or territory of residence of children registered in the CCTB program. Therefore, monthly, quarterly and annual migration can be estimated by comparing two microdata files.

The estimation of interprovincial migration of 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 who are registered to the CCTB program, for each month and province and territory. The second factor, which corrects for any biases resulting from the differential propensity to migrate between children who are registered to the CCTB program and all children, is obtained by comparing the interprovincial out-migration rates of children who are registered to the CCTB program with the out-migration rates for all children (aged 0 to 17). The adjustment factor is calculated for each province and territory and is based on the two last available years of T1FF.

A matrix is used to represent interprovincial migration between 13 provinces and territories of origin (denoted by j) and 13 destination provinces and territories (denoted by k). Each item (j,k) of the matrix represents both the number of out-migrants from province or territory of origin j and the number of in-migrants to destination province or territory k.

The formula used to estimate each item (j,k) of the matrix for the interprovincial migration of children is as follows:

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

where

${}_{j,k}M^{0-17}$ = number of children migrating from province or territory of origin j to destination province or territory k;

${}^{CCTB}_{j,k}M^{0-17}$ = number of children in the CCTB database migrating from province or territory of origin j to destination province or territory k;

${}^{CCTB}_{j}\mathfrak{R}$ = coverage rate of the CCTB program for the province or territory of origin j.

${}_{j}G$ = adjustment factor for differential migration propensity among children for the province or territory of origin j.

The equation above includes the adjustments for incomplete coverage of CCTB-children who are registered to the CCTB program (equation 7.2) and for the differences in the propensities of this group to migrate as compared to all children (equation 7.3).

The coverage rates are calculated as follows:

For each province or territory,

Equation 7.2: $$ {}^{CCTB}_{j}\mathfrak{R} = \left[\frac{{}^{CCTB}_{j}P^{0-17}}{{}^{Dem}_{j}P^{0-17}} \right] $$

where

${}^{CCTB}_{j}P^{0-17}$ = number of children in province or territory j on the basis of CCTB data; and

${}^{Dem}_{j}P^{0-17}$ = number of children in province or territory j according to Demography Division estimates.

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

For each province and territory:

Equation 7.3 : $$ {}_{j}G = \left[\frac{{}^{TAX}_{j}Rate\_M^{0-17}}{{}^{CCTB}_{j}Rate\_M^{0-17}} \right] = \frac{\left[\frac{{}^{TAX}_{j}M^{0-17}}{{}^{TAX}_{j}P^{0-17}} \right]}{\left[\frac{{}^{CCTB}_{j}M^{0-17}}{{}^{CCTB}_{j}P^{0-17}} \right]} $$

where

${}^{TAX}_{j}Rate\_M^{0-17}$ = interprovincial out-migration rate of all children from province or territory of origin j, according to income tax data;

${}^{CCTB}_{j}Rate\_M^{0-17}$ = interprovincial out-migration rate of children who are registered to the CCTB program from province or territory of origin j, as modeled from tax data (T1FF);

${}^{TAX}_{j}M^{0-17}$ = number of total out-migrant children from province or territory of origin j, according to income tax data;

${}^{TAX}_{j}P^{0-17}$ = estimate of children in province or territory j based on income tax files;

${}^{CCTB}_{j}M^{0-17}$ = number of out-migrant children who are registered to the CCTB program from province or territory j, according to tax data;

${}^{CCTB}_{j}P^{0-17}$ = Children who are registered to the CCTB program in the population of the income tax files provided by T1FF, by province or territory j.

The ${}_{j}G$ factor will equal one if the migration rate of children who are registered to the CCTB program and the migration rate of all children are identical. A value of ${}_{j}G$ that is greater (or less) than one indicates that the propensity to migrate of children who are registered to the CCTB program is less (or greater) than that of all children.

Interprovincial migration of adults

The second factor for adjusting the estimates produced using CCTB data is used in the calculation of the number of adult migrants. Estimates of adults migrating across provincial or territorial boundaries are obtained by calculating modelled migration ratios of adults to children, called F factors. F factors, which are calculated using the most recent tax data, are the ratio of the adult migration rate to the child migration rate, by province or territory of origin and by destination province or territory, as follows:

For each province and territory of origin and each destination province and territory,

Equation 7.4: $$ {}_{j,k}F_{t-3,t} = \left[\frac{{}_{j,k}Avg\_Rate\_M^{18+}}{{}_{j,k}Avg\_Rate\_M^{0-17}} \right] = \frac{\left[\frac{{}^{TAX}_{j,k}M^{18+}_{t-3,t-2} + {}^{TAX}_{j,k}M^{18+}_{t-2,t-1} + {}^{TAX}_{j,k}M^{18+}_{t-1,t}}{{}^{Dem}_{j}P^{18+}_{t-2}} \right]}{\left[\frac{{}^{TAX}_{j,k}M^{0-17}_{t-3,t-2} + {}^{TAX}_{j,k}M^{0-17}_{t-2,t-1} + {}^{TAX}_{j,k}M^{0-17}_{t-1,t}}{{}^{Dem}_{j}P^{0-17}_{t-2}} \right]} $$

where

$ {}_{j,k}F_{t-3,t} $ = average estimation factor for adults calculated over three migration periods by province or territory of origin j and destination province or territory k;

$ {}_{j,k}Avg_Rate_M^{18+} $ = average rate of interprovincial migration of adults over three migration periods by province or territory of origin j and destination province or territory k;

$ {}_{j,k}Avg_Rate_M^{0-17} $ = average rate of interprovincial migration of children over three migration periods by province or territory of origin j and destination province or territory k;

$ {}^{TAX}_{j,k}M^{18+} $ = number of adult migrants by province or territory of origin j and destination province or territory k on the basis of income tax data for period (t-3, t-2), (t-2, t-1) and (t-1, t);

$ {}^{Dem}_{j}P^{18+}_{t-2} $ = number of adults for the province or territory of origin j at time (t 2) according to Demography Division estimates;

$ {}^{TAX}_{j,k}M^{0-17} $ = number of child migrants by province or territory of origin j and destination province or territory k on the basis of income tax data for periods (t-3, t-2), (t-2, t-1) and (t-1, t); and

$ {}^{Dem}_{j}P^{0-17}_{t-2} $ = number of children for the province or territory of origin j at time (t-2) according to Demography Division estimates.

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

For each province and territory,

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

where

$ {}_{j,k}M^{18+} $ = estimated number of adults migrating from province or territory of origin j to destination province or territory k;

$ \frac{{}_{j,k}M^{0-17}}{{}^{Dem}_{j}P^{0-17}} $ = rate of migration of children (estimates by province or territory of origin j and destination province or territory k on the basis of adjusted CCTB data and Demography Division population estimates); and

$ {}^{Dem}_{j}P^{18+} $ = number of adults in province or territory of origin j at the beginning of the period according to Demography Division population estimates.

Finally, after estimates for the total numbers of interprovincial child and adult migrants have been separately calculated by province or territory of origin j and destination province or territory k, interprovincial migration estimates are obtained by adding the number of interprovincial child and adult migrants as follows:

For each province and territory,

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

7.2.2 Monthly, quarterly and annual estimates

The monthly, quarterly and annual number of migrants can be calculated using monthly CCTB microdata. However, adjustments must be made to the numbers of migrants to align the net interprovincial migration of the different measurement periods for which estimates are derived.

Aligning monthly net numbers with quarterly net numbers

Each quarter, monthly data must be adjusted to align the monthly net numbers with the quarterly net numbers.

First, the difference between the net values of the sum of the three monthly matrices and those of the quarterly matrix is calculated for each combination of origin and destination:

Equation 7.7: $${}^{P}_{j,k}\Delta_{t,t+2} = \left({}^{P}_{k,j}SM_{t,t+2} - {}^{P}_{j,k}SM_{t,t+2}\right) - \left({}^{P}_{k,j}Q_{t,t+2} - {}^{P}_{j,k}Q_{t,t+2}\right)$$

where

$${}^{P}_{j,k}\Delta_{t,t+2}$$ = difference in number of migrants between the sum of the preliminary (P) monthly matrices and the preliminary (P) quarterly matrix for the province or territory of origin j, and the destination province or territory k;

$${}^{P}_{j,k}SM_{t,t+2}$$ = sum of the preliminary (P) monthly number of migrants from month t to month t+2 for the province or territory of origin j, and the destination province or territory k; and

$${}^{P}_{j,k}Q_{t,t+2}$$ = preliminary (P) quarterly number of migrants for the quarter covering months t to t+2 for the province or territory of origin j, and the destination province or territory k.

These differences are used to change the number of migrants to create a new matrix in which the net numbers match those of the quarterly matrix.

Each item (j,k) of the adjusted matrix is obtained as follows:

Equation 7.8: $${}^{P}_{j,k}SM^{*}_{t,t+2} = {}^{P}_{j,k}SM_{t,t+2} + {}^{P}_{j,k}\Delta_{t,t+2} \times \frac{{}^{P}_{j,k}SM_{t,t+2}}{{}^{P}_{j,k}SM_{t,t+2} + {}^{P}_{k,j}SM_{t,t+2}}$$

where

$${}^{P}_{j,k}SM^{*}_{t,t+2}$$ = sum of the preliminary (P) monthly number of migrants from month t to month t+2 adjusted so that net numbers match those of the quarterly matrix for the province or territory of origin j, and the destination province or territory k.

Various adjustments are made to matrix $${}^{P}SM^{*}_{t,t+2}$$ to avoid negative values or division by zero.

Lastly, for each of the three months in the quarter, the adjusted monthly interprovincial migration matrix is derived using the matrix resulting from the equation above. Each matrix item (j,k) is adjusted as follows:

Equation 7.9: $${}^{P}_{j,k}M^{*}_{t} = {}^{P}_{j,k}M_{t} \times \frac{{}^{P}_{j,k}SM^{*}_{t,t+2}}{{}^{P}_{j,k}SM_{t,t+2}}$$

where

$${}^{P}_{j,k}M^{*}_{t}$$ = preliminary (P) monthly number of migrants for month t adjusted so that net numbers match quarterly net numbers for the province or territory of origin j, and the destination province or territory k; and

$${}^{P}_{j,k}M_{t}$$ = preliminary (P) monthly number of migrants for month t for the province or territory of origin j, and the destination province or territory k.

This matrix contains the preliminary monthly numbers of migrants adjusted so that net numbers match the quarterly net interprovincial numbers.

Aligning monthly net numbers with annual net numbers

Each year, the monthly interprovincial migration estimates must be adjusted to align the monthly net numbers with the net numbers of the annual matrix.

The adjustment involves calculating the difference between the net numbers of the sum of the 12 monthly matrices ${}^{P}M_{t}^{*}$ and the net numbers of the annual matrix for each combination of origin and destination:

Equation 7.10: $${}^{P}_{j,k}\Delta_{t,t+11} = \left({}^{P}_{k,j}SM^{*}_{t,t+11} - {}^{P}_{j,k}SM^{*}_{t,t+11}\right) - \left({}^{P}_{k,j}A_{t,t+11} - {}^{P}_{j,k}A_{t,t+11}\right)$$

where

${}^{P}_{j,k}\Delta_{t,t+11}$ = difference in number of migrants between the sum of the preliminary (P) monthly matrices adjusted for the quarters and the preliminary (P) annual migration matrix for the province or territory of origin j, and the destination province or territory k;

${}^{P}_{j,k}SM^{*}_{t,t+11}$ = sum of the preliminary (P) monthly number of migrants from month t to month t+11 adjusted for the quarters for the province or territory of origin j and the destination province or territory k; and

${}^{P}_{j,k}A_{t,t+11}$ = preliminary (P) annual number of migrants for the year covering months t to t+11 for the province or territory of origin j, and the destination province or territory k.

These differences are then used to change the number of migrants to create a new matrix in which the monthly net numbers match the annual net numbers.

Each item (j,k) of this matrix is calculated as follows:

Equation 7.11:

$${}^{P}_{j,k}SM^{**}_{t,t+11} = {}^{P}_{j,k}SM^{*}_{t,t+11} + {}^{P}_{j,k}\Delta_{t,t+11} \times \frac{{}^{P}_{j,k}SM^{*}_{t,t+11}}{{}^{P}_{j,k}SM^{*}_{t,t+11} + {}^{P}_{k,j}SM^{*}_{t,t+11}}$$

where

${}^{P}_{j,k}SM^{**}_{t,t+11}$ = sum of the preliminary (P) monthly number of migrants from month t to month t+11 for the province or territory of origin j, and the destination province or territory k the net numbers of which match the preliminary annual net numbers.

Adjustments are made to matrix ${}^{P}SM^{**}_{t,t+11}$ to avoid negative values or division by zero.

Lastly, for each of the 12 months, the adjusted monthly interprovincial migration matrix is derived by obtaining each item (j,k) as follows:

Equation 7.12: $${}^{P}_{j,k}M^{**}_{t} = {}^{P}_{j,k}M^{*}_{t} \times \frac{{}^{P}_{j,k}SM^{**}_{t,t+11}}{{}^{P}_{j,k}SM^{*}_{t,t+11}}$$

where

${}^{P}_{j,k}M^{**}_{t}$ = preliminary (P) monthly number of migrants for month t adjusted so that net numbers match annual net numbers for the province or territory of origin j, and the destination province or territory k.

This matrix contains the preliminary monthly numbers of migrants adjusted so that the net numbers match the annual net interprovincial numbers.

Aligning quarterly net numbers with annual net numbers

Each year, when the annual matrix is created, the quarterly net numbers must also be aligned with the annual net numbers.

The adjustment involves calculating the difference between the net numbers of the quarterly matrix and those of the sum of the matrices for the three months of the quarter. Note that net interprovincial migration in the preliminary monthly matrices has already been aligned with annual net numbers.

Equation 7.13: $${}^{P}_{j,k}\Delta^{*}_{t,t+2} = \left({}^{P}_{k,j}Q_{t,t+2} - {}^{P}_{j,k}Q_{t,t+2}\right) - \left({}^{P}_{k,j}SM^{**}_{t,t+2} - {}^{P}_{j,k}SM^{**}_{t,t+2}\right)$$

where

$${}^{P}_{j,k}\Delta^{*}_{t,t+2}$$ = difference in number of migrants between the preliminary (P) quarterly migration matrix and the sum of the preliminary (P) monthly matrices adjusted for annual net numbers for the province or territory of origin j, and the destination province or territory k.

The differences are then used to change the quarterly number of migrants so that the net numbers of the quarterly matrix match those of the sum of the net numbers of the monthly matrices for the three corresponding months.

Each item (j,k) of the adjusted matrix is obtained as follows:

Equation 7.14: $${}^{P}_{j,k}T^{*}_{t,t+2} = {}^{P}_{j,k}T_{t,t+2} + {}^{P}_{j,k}\Delta^{*}_{t,t+2} \times \frac{{}^{P}_{j,k}T_{t,t+2}}{{}^{P}_{j,k}T_{t,t+2} + {}^{P}_{k,j}T_{t,t+2}}$$

where

$${}^{P}_{j,k}Q^{*}_{t,t+2}$$ = preliminary (P) quarterly number of migrants covering months t to t+2 adjusted so that net numbers match annual net numbers for the province or territory of origin j, and the destination province or territory k.

Adjustments are made to matrix $${}^{P}_{j,k}Q^{*}_{t,t+2}$$ to avoid negative values or division by zero.

This matrix contains the preliminary quarterly number of migrants adjusted so that the net numbers match the annual net interprovincial numbers.

7.3 Final interprovincial migration estimates, Canada, provinces and territories

7.3.1 Estimation methodology

Final estimates of interprovincial migration are produced using personal income tax data. By comparing the place of residence at the time of filing, it is possible to identify for those 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 steps:[39]

1. geocoding 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 T1FF.

The four steps of the estimation methodology are described below.

Step 1 – Geocoding

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 (SGC) codes used by the census. About 99% of all tax records are filed with a postal code, or are assigned

39. For a detailed description of the methodology, see Migration Estimates. User Guide, Demography Division, Statistics Canada. Ottawa. Internal document. 2010.

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 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 or 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 Statistics Canada's vital statistics birth database, and in addition a historical file of imputed children is created from the previous year's tax file.[40]

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.

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 to female ratio based on information from the most recent census.

Step 3 – Identifying migrant tax filers

After assigning the geographic codes and non-filing dependants to each tax filer, records for two consecutive years are matched by social insurance number. Only records of tax filers who are present in both years are retained. Migrant tax filers are identified by comparing current and previous geography codes (CDs or CMAs). Thus, it is only possible to determine migrant status for those who file 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. However, in the Demographic Estimates Program the given estimates cover the period from July to June. 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 factor

The final step in the estimation process is an adjustment for coverage, done at the CD to CMA level for gender and specific age groups. This adjustment is required to estimate those migrants who did not file two consecutive personal income tax returns, which involves the calculation of coverage adjustment factors (or inflation factors). Population estimates by CD to ratios CMA are used as the denominator to the same population for whom two consecutive income tax returns could be linked to create coverage ratios.

For migration estimates prior and up to 2000/2001, provincial adjustment ratios were substituted in place of the CD to CMA ratio in the few cases where coverage was abnormally high or low (outlier values in the coverage ratios). Beginning with 2001/2002 migration data, high and low coverage cases were identified with a different methodology and a Canadian adjustment ratio was used in place of the CD to CMA ratio for these outlier coverage ratios. After a review and analysis of these methods, it was determined that cases where the coverage ratios were deemed outlier, represented small domains having little impact on the measure of migration, moreover, it was known that replacing the coverage ratios with a national ratio had the effect of reducing the true measure of migration flows. Consequently, beginning in 2006/2007, the identification and treatment of high and low coverage cases was discontinued and the coverage was simply adjusted to the Demography Division data at the CD level by age and sex group. The same adjustment is also done for CMAs.

40. The Historical Family File, updated annually by Income Statistics Division of Statistics Canada, consists of information of known family relationships obtained from tax returns.

The adjustment ratios are applied to the counts of interprovincial 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 rate of migration as those for whom it can be measured.

7.3.2 Monthly, quarterly and annual estimates

Final annual estimates are obtained directly from T1FF data. However, since these data are annual, models must be used to calculate monthly and quarterly estimates. As well, these estimates must be adjusted to align the net interprovincial migration of the different measurement periods for which estimates are derived. These adjustments are similar to those for the preliminary estimates. However, for clarity, they are reproduced here.

As is the case with preliminary estimates, matrix representations are used for final estimates. A matrix is used to represent interprovincial migration between 13 provinces and territories of origin (denoted by j) and 13 destination provinces and territories (denoted by k). Each item (j,k) of the matrix represents both the number of out-migrants from province or territory of origin j and the number of in-migrants to destination province or territory k.

Monthly estimates

Monthly interprovincial migration estimates are derived as follows.

First, a matrix containing the sum of the monthly number of migrants in the 12 months of the year is created using the annual matrix of the T1FF and the preliminary estimates derived from CCTB data.

For each province or territory of origin j and destination province or territory k, item (j,k) of the matrix is calculated as follows:

Equation 7.15: $$ {}^{D}_{j,k}SM_{t,t+11} = {}^{P}_{j,k}SM^{**}_{t,t+11} \times \frac{{}^{D}_{j,k}A_{t,t+11}}{{}^{P}_{j,k}A_{t,t+11}} $$

where

$ {}^{D}_{j,k}SM_{t,t+11} $ = sum of the final (D) monthly number of migrants from month t to month t+11 for the province or territory of origin j, and the destination province or territory k;

$ {}^{P}_{j,k}SM^{**}_{t,t+11} $ = sum of the preliminary (P) monthly number of migrants from month t to month t+11 for the province or territory of origin j, and the destination province or territory k aligned with preliminary annual net numbers;

$ {}^{D}_{j,k}A_{t,t+11} $ = final (D) annual number of migrants for the province or territory of origin j, and the destination province or territory k; and

$ {}^{P}_{j,k}A_{t,t+11} $ = preliminary (P) annual number of migrants for the province or territory of origin j, and the destination province or territory k.

An adjustment is then applied to the matrix so that the sum of monthly net interprovincial migration matches that of the annual matrix.

The adjustment involves calculating the difference between the net numbers of the two matrices for each combination of origin and destination:

Equation 7.16: $$ {}^{D}_{j,k}\Delta_{t,t+11} = \left({}^{D}_{k,j}SM_{t,t+11} - {}^{D}_{j,k}SM_{t,t+11}\right) - \left({}^{D}_{k,j}A_{t,t+11} - {}^{D}_{j,k}A_{t,t+11}\right) $$

where

$ {}^{D}_{j,k}\Delta_{t,t+11} $ = difference in number of migrants between the final (D) matrix resulting from the sum of the monthly number of migrants from month t to month t+11 and the final annual matrix for the province or territory of origin j, and the destination province or territory k.

The differences calculated above are then used to change the number of migrants to create a new matrix in which the net numbers match those of the annual migration matrix.

Each item (j,k) of the adjusted matrix is obtained as follows:

Equation 7.17: $${}^{D}_{j,k}SM^{*}_{t,t+11} = {}^{D}_{j,k}SM_{t,t+11} + {}^{D}_{j,k}\Delta_{t,t+11} \times \frac{{}^{D}_{j,k}SM_{t,t+11}}{{}^{D}_{j,k}SM_{t,t+11} + {}^{D}_{k,j}SM_{t,t+11}}$$

where

$${}^{D}_{j,k}SM^{*}_{t,t+11}$$ = sum of the final (D) monthly number of migrants adjusted so that net numbers match those of the annual matrix for the province or territory of origin j, and the destination province or territory k.

Various adjustments are made to matrix $${}^{D}SM^{*}_{t,t+11}$$ to avoid negative values or division by zero.

Lastly, for each of the 12 months, the adjusted monthly interprovincial migration matrix is derived using the matrix resulting from equation 7.17 and preliminary monthly estimates derived from the CCTB.

Equation 7.18: $${}^{D}_{j,k}M^{*}_{t} = {}^{P}_{j,k}M^{**}_{t} \times \frac{{}^{D}_{j,k}SM^{*}_{t,t+11}}{{}^{P}_{j,k}SM^{**}_{t,t+11}}$$

where

$${}^{D}_{j,k}M^{*}_{t}$$ = final (D) monthly number of migrants in month t adjusted so that net numbers match the annual net numbers for the province or territory of origin j, and the destination province or territory k; and

$${}^{P}_{j,k}M^{**}_{t}$$ = preliminary (P) monthly number of migrants in month t adjusted so that preliminary net numbers match the preliminary annual net numbers for the province or territory of origin j, and the destination province or territory k.

This matrix contains the final monthly numbers of migrants adjusted so that the sum of the net numbers matches that of the annual matrix.

Quarterly estimates

When monthly interprovincial migration estimates are produced, quarterly migration estimates can be derived.

First, final quarterly matrices must be derived using information from preliminary matrices and final monthly matrices for the three months in each quarter produced above.

For each province or territory of origin j and destination province or territory k, each item (j,k) of the quarterly migration matrix is calculated as follows:

Equation 7.19: $${}^{D}_{j,k}Q_{t,t+2} = {}^{P}_{j,k}Q^{*}_{t,t+2} \times \frac{{}^{D}_{j,k}SM^{*}_{t,t+2}}{{}^{P}_{j,k}SM^{**}_{t,t+2}}$$

where

$${}^{D}_{j,k}Q_{t,t+2}$$ = final (D) quarterly number of migrants from month t to month t+2 for the province or territory of origin j and the destination province or territory k; and

$${}^{P}_{j,k}Q^{*}_{t,t+2}$$ = preliminary (P) quarterly number of migrants from month t to month t+2 aligned with preliminary annual net numbers for the province or territory of origin j, and the destination province or territory k.

An adjustment is then applied to quarterly matrix $${}^{D}Q_{t,t+2}$$ to align net numbers with those of the sum of the three corresponding monthly matrices.

The adjustment involves calculating the difference between the net numbers of the quarterly matrix and those of the sum of the monthly matrices for the three months of the quarter. Note that net interprovincial migration in the final monthly matrices has already been aligned with the annual net numbers.

Equation 7.20: $${}^{D}_{j,k}\Delta_{t,t+2} = ({}^{D}_{k,j}Q_{t,t+2} - {}^{D}_{j,k}Q_{t,t+2}) - ({}^{D}_{k,j}SM^{*}_{t,t+2} - {}^{D}_{j,k}SM^{*}_{t,t+2})$$

The differences are then used to change the number of migrants in the quarterly matrix to align the quarterly net numbers with those of the sum of the three monthly matrices for the three months of the quarter.

Each item (j,k) of the adjusted matrix is obtained as follows:

Equation 7.21: $${}^{D}_{j,k}Q^{*}_{t,t+2} = {}^{D}_{j,k}Q_{t,t+2} + {}^{D}_{j,k}\Delta_{t,t+2} \times \frac{{}^{D}_{j,k}Q_{t,t+2}}{{}^{D}_{j,k}Q_{t,t+2} + {}^{D}_{k,j}Q_{t,t+2}}$$

where

$${}^{D}_{j,k}Q^{*}_{t,t+2}$$ = final (D) quarterly number of migrants adjusted so that net numbers match those of the sum of the three corresponding monthly matrices for the province or territory of origin j, and the destination province or territory k.

Various adjustments are made to matrix $${}^{D}_{j,k}Q^{*}_{t,t+2}$$ to avoid negative values or division by zero.

This matrix contains the final quarterly numbers of migrants adjusted so that the net numbers match those of the sum of the three corresponding monthly matrices.

7.4 Interprovincial migration by age and sex

Final interprovincial migration estimates are derived from the T1FF file consisting solely of five broad age groups (under 18, 18 to 24, 25 to 44, 45 to 64, and 65 years and over) by sex, as well as the origin and destination province or territory.

To estimate the number of interprovincial migrants by age and sex, the following steps are followed:

i. calculate migration by sex and five-year age cohorts using the number of migrants by province of origin and destination obtained from the one-year ago mobility question from the most recent census;

ii. split the migration cohort into single-year of age using Sprague's multipliers;

iii. calculate a percentage distribution by single year of age for each broad age-sex group (under 18, 18 to 24, 25 to 44, 45 to 64, and 65 years and over), using the data obtained in step (ii);

iv. apply the distribution obtained in step (iii) (which remains constant for the five-year postcensal estimates) to the interprovincial migration estimates by broad age group and sex, in order to disaggregate them into a distribution by single-year of age;

v. finally, to eliminate inconsistencies from one age to another, the number of migrants by single-year of age, is aggregated by five-year age groups, to which Sprague's multipliers are again applied.

To produce the preliminary interprovincial migration estimates by age and sex at the provincial and territorial level, the most recent age and sex distribution of final estimates is applied.

Chapter 8

Subprovincial population estimates

In addition to estimates at the national and provincial and territorial levels, population estimates are produced for subprovincial regions. Annual population estimates by age and sex are released for census divisions (CDs), census metropolitan areas (CMAs) and economic regions (ERs). Custom requests for population estimates at other geographical levels (e.g., census subdivisions (CSDs), census agglomerations (CAs) and health regions) are also available. Annual population estimates for the subprovincial regions are based on the most recent Standard Geographical Classification (SGC) from 2001 on.

This chapter discusses the estimation methods used to produce postcensal and intercensal population estimates for CDs, CMAs and ERs. The methods used to produce population estimates at the census subdivision level are also presented at the end of this chapter.

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 province or territory level and the municipality (i.e., census subdivision).[41] 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, Northwest Territories and Nunavut, provincial or territorial law does not provide for these administrative geographic areas. Therefore, Statistics Canada, in cooperation with these provinces and territories, has created equivalent areas called census divisions for the purpose of disseminating statistical data. In Yukon, the census division is equivalent to the entire territory.

A census metropolitan area (CMA) is formed by one or more adjacent municipalities centred on a population centre (known as the core). A CMA must have a total population of at least 100,000 of which 50,000 or more must live in the core. To be included in the CMA, other adjacent municipalities must have a high degree of integration with the core, as measured by commuting flows derived from previous census place of work data.

Once an area becomes a CMA, it is retained as a CMA even if its total population declines below 100,000 or the population of its core falls below 50,000. Small population centres with a population count of less than 10,000 are called fringe. All areas inside the CMA that are not population centres are rural areas.

An economic region (ER) is a grouping of complete CDs (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 and territories, economic regions are created by agreement between Statistics Canada and the province or territory concerned. Prince Edward Island and the three territories each consist of one ER. In Ontario, there is one exception where the ER 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 regions

8.2.1 Postcensal population estimates for CMAs and CDs

The component method is used to produce population estimates for CMAs and CDs. To ensure concordance between the subprovincial and provincial and territorial population estimates by age and sex, two-way raking is used.

The following formula is used to produce the total CMA and CD population estimates:

For each subprovincial region:

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

Equation 8.1:

$$P_{(t+1)} = P_{(t)} + B_{(t,t+1)} - D_{(t,t+1)} + I_{(t,t+1)} - \left(E_{(t,t+1)} + \Delta TE_{(t,t+1)}\right) + RE_{(t,t+1)} + \Delta NPR_{(t,t+1)} + \Delta Ninter_{(t,t+1)} + \Delta Ninfra_{(t,t+1)}$$

where, for each subprovincial region

$(t, t+1)$ = interval between times t and t+1;

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

$P_{(t)}$ = base population at time t (census counts adjusted for census net undercoverage or the most recent estimate);

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

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

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

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

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

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

$\Delta NPR_{(t,t+1)}$ = net non-permanent residents;

$\Delta Ninter_{(t,t+1)}$ = net interprovincial migration;

$\Delta Ninfra_{(t,t+1)}$ = net intraprovincial migration.

Subprovincial or intraprovincial migration (migration within one same province or territory and between subprovincial regions) is a necessary additional component to estimate migration at the subprovincial level.

8.2.2 Postcensal population estimates for economic regions (ERs)

A different method is used to produce population estimates for economic regions (ERs). In this case the census division's (CD) aggregate method is used. First, the ERs are defined in terms of CDs using the most recent Standard Geographical Classification (SGC) specifications. When the geographic delineation of the CDs and ERs are the same, 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 CD does not match that of the ER, i.e., when a CD is in more than one ER, distribution of the CD's demographic components are allocated on the basis of its demographic weight in each ER in question. The proportions are referred to as conversion factors. They are calculated using the most recent census counts.

Thus, demographic components (births, deaths and migration) initially measured at the CD level can be allocated to each ER. Using the census division's aggregate method by the ERs' geographic delineation, the population and demographic components of ERs can be estimated.

However, the census division's aggregate method cannot be used to estimate the number of intraprovincial in-migrants and out-migrants, since it overestimates those figures. In-migrants to a given CD from another CD in the same ER should not be counted since the migration occurred within the ER's boundaries. These are false

in-migrants. The same is true for out-migrants from one CD to another CD in the same ER: they are false out-migrants. However, the net intraprovincial migration calculated with the CD aggregate method is correct because the false in-migrants and out-migrants cancel each other out. As a result, only the net intraprovincial migration of ERs can be estimated accurately using the CD aggregate method. This is why the estimates for intraprovincial in-migrants and out-migrants are not available at the ER level.

8.2.3 Subprovincial postcensal population estimates by age and sex

The component method is used to produce postcensal population estimates by age and sex for CMAs and CDs. The method is applied to each age-sex cohort in the base population. Chapter 9 describes the application of the cohort component method in detail.

The component method formulas for estimating the population of CMAs and CDs by age and sex are as follows:

For age 0:

Equation 8.2:

$$P^{0}_{(t+1)} = B_{(t,\,t+1)} - D^{-1}_{(t,\,t+1)} + I^{-1}_{(t,\,t+1)} - \left(E^{-1}_{(t,\,t+1)} + \Delta TE^{-1}_{(t,\,t+1)}\right) + RE^{-1}_{(t,\,t+1)} + NPR^{0}_{(t,\,t+1)} + \Delta Ninter^{-1}_{(t,\,t+1)} + \Delta Ninfra^{-1}_{(t,\,t+1)}$$

For ages 1 to 89:

Equation 8.3:

$$P^{a+1}_{(t+1)} = P^{a}_{(t)} - D^{a}_{(t,\,t+1)} + I^{a}_{(t,\,t+1)} - \left(E^{a}_{(t,\,t+1)} + \Delta TE^{a}_{(t,\,t+1)}\right) + RE^{a}_{(t,\,t+1)} - NPR^{a}_{(t)} + NPR^{a+1}_{(t,\,t+1)} + \Delta Ninter^{a}_{(t,\,t+1)} + \Delta Ninfra^{a}_{(t,\,t+1)}$$

For age group 90 and over:

Equation 8.4:

$$P^{90+}_{(t+1)} = P^{89+}_{(t)} - D^{89+}_{(t,\,t+1)} + I^{89+}_{(t,\,t+1)} - \left(E^{89+}_{(t,\,t+1)} + \Delta TE^{89+}_{(t,\,t+1)}\right) + RE^{89+}_{(t,\,t+1)} - NPR^{89+}_{(t)} + NPR^{90+}_{(t+1)} + \Delta Ninter^{89+}_{(t,\,t+1)} + \Delta Ninfra^{89+}_{(t,\,t+1)}$$

where, for each subprovincial region

$(t, t+1)$ = interval between times t and t+1;

$P_{(t+1)}$ = population estimates at time t+1;

$P_{(t)}$ = base population at time t (census counts adjusted for net census undercoverage or the most recent estimate);

B = number of births;

D = number of deaths;

I = number of immigrants;

E = number of emigrants;

ΔTE = net temporary emigration;

RE = number of returning emigrants;

NPR = number of non-permanent residents;

$\Delta Ninter$ = net interprovincial migration;

$\Delta Ninfra$ = net intraprovincial migration;

To ensure concordance between the subprovincial estimates and the provincial and territorial estimates by age and sex, two-way raking is used.

Special treatment for preliminary postcensal estimates for Quebec and British Columbia

A different method is used to calculate preliminary postcensal population estimates for census divisions (CDs) and census metropolitan areas (CMAs) in Quebec. The population estimates by age and sex produced by the *Institut de la statistique du Québec* (ISQ) are used. These population estimates are based on data from the *Fichier d'inscription des personnes assurées* (FIPA), the insured persons register, from the *Régie de l'assurance-maladie du Québec* (RAMQ).

For British Columbia, preliminary postcensal estimates at the CMA and CD levels are calculated by applying the total population growth rates provided by BC Stats, British Columbia's statistical agency, to the previous year's estimates produced by the Demography Division. The total preliminary postcensal estimates are then distributed by age and sex using the Demography Division's component method. The British Columbia population estimates used to calculate the rates are produced using a regression model based on data from residential Hydro services and Ministry of Health Client Registry data as symptomatic indicators.

To ensure concordance between the subprovincial estimates and the provincial totals by age and sex, two-way raking is used.

8.2.4 Estimate levels

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

For the subprovincial regions in 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 deemed preliminary postcensal (PP). When the components are all final, the estimate is deemed final postcensal (PD). Any other combination of levels is considered updated postcensal (PR).

8.2.5 Base population and components of population growth

Base population

A full description of the methodology used to calculate the postcensal base population is described in Chapter 2. In the Demographic Estimates Program, the base populations for subprovincial regions are derived from the five-year censuses between 2001 and 2011. Population counts at the provincial, territorial and subprovincial levels are subject to the same adjustment procedures outlined in Chapter 2, unless otherwise noted. To estimate census net undercoverage (CNU) at the subprovincial level, provincial and territorial CNU rates by age and sex are applied to census subdivisions (CSDs), which are aggregated to create the base population of higher subprovincial levels (census metropolitan areas (CMAs) and census divisions (CDs) in the province).

Prior to generating the demographic estimates for each component, two-way raking is used to ensure concordance between the subprovincial and the provincial or territorial totals by age and sex.

Two-way raking involves simultaneously adjusting the differences by assuming consistency between the following:

1. the sum of the population of subprovincial regions by province or territory and the total provincial and territorial population, and
2. the sum of the population of subprovincial regions by age and sex and the provincial and 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 of the Standard Geographical Classification (SGC), as defined in the most recent census. This ensures a stable geographical universe for the whole reference period.

Births and deaths

The numbers of births and deaths at the census division (CD) and census metropolitan area (CMA) levels are taken directly from the vital statistics database of Statistics Canada's Health Statistics Division. For CMAs, births and deaths have been calculated using this data source since 2007/2008.[42]

Births and deaths estimates are categorized as final when they are directly taken from the Health Statistics Division's vital statistics. They are then adjusted to the provincial and territorial totals using two-way raking to ensure their concordance.

When no births or deaths data are available, preliminary provincial or territorial estimates are broken down, using the most recent known subprovincial distribution derived from Health Statistics Division's vital statistics, to produce estimates by region. In that case, the births and deaths estimates are categorized as preliminary. They are then adjusted to the provincial and territorial totals using two-way raking to ensure their consistency.

Immigration

Since we do not use subprovincial immigration data from Citizenship and Immigration Canada (CIC), the most recent known subprovincial distribution derived from the T1FF[43] is used to produce immigrant estimates by subprovincial region. Because the data are available only by broad age groups (0-17, 18-24, 25-44, 45-64, 65+), they are broken down by age and sex based on the distribution from the most recent census or NHS (starting in 2011). The distribution stems from the NHS mobility question on place of residence one year ago. Since 2011/2012, NHS distributions have been modelled to minimize the impact of outliers found in some subprovincial regions, mostly for smaller geographies. To ensure their consistency, subprovincial estimates are then adjusted to the provincial and territorial totals using two-way raking.

The difference between preliminary and final estimates lies in the timeliness of the sources used to estimate this component. Since the subprovincial estimates of immigrants are adjusted to provincial and territorial estimates, the level of subprovincial estimates will be the same. Immigration estimates are preliminary the first year and final the following year.

Net non-permanent residents (NPRs)

At the subprovincial level, there are no reliable administrative data sources available to directly estimate net NPRs. To compensate for this lack of data, the provincial and territorial NPR estimates by age and sex are broken down by subprovincial region based on the distribution from the most recent census or NHS (starting in 2011). Since 2011/2012, NHS distributions have been modelled to minimize the impact of outliers found in some subprovincial regions, mostly for smaller geographies. To ensure their consistency, subprovincial estimates are then adjusted to the provincial and territorial totals using two-way raking.

For the 2005/2006 and 2010/2011 years, the net NPRs are calculated using two different distributions—the 2001 and 2006 censuses for the year 2005/2006, and the 2006 Census as well as the 2011 NHS for the year 2010/2011. This approach assumes that the two distributions are similar. If the two distributions vary by the regional breakdown of NPRs, the net NPRs for 2005/2006 and 2010/2011 will absorb all the changes attributable to the difference between the two distributions that were used. For this reason, the net NPRs for 2005/2006 and 2010/2011 should not be compared with the rest of the historical series.

Since the subprovincial estimates of the net number of NPRs are adjusted to provincial and territorial estimates, the level of the subprovincial estimates will be the same. NPR estimates are preliminary the first year and updated the following year. They become final two to three years after the reference year, when all other components are also final.

42. To produce birth and death estimates for previous periods, the geographical conversion factor method was used. According to this method, each CMA is first defined in terms of CDs (complete or partial) using the Standard Geographical Classification (SGC) specifications. These proportions are also known as conversion factors and are calculated using the most recent census counts. With these factors and census division's birth and death estimates, it is possible to allocate, to each CMA, the births and deaths initially measured at the CD level.
43. The T1 Family File (T1FF) is drawn from the Canada Revenue Agency (CRA) T1 file by the Statistics Canada's Income Statistics Division. In this document, the acronym T1FF refers to this file.

Emigration

As with immigrants, the number of emigrants at the subprovincial level is derived from the T1FF. Because the estimates are available only by broad age groups (0-17, 18-24, 25-44, 45-64, 65+), they are broken down by age and sex based on the provincial or territorial distribution. They are then adjusted to the provincial and territorial totals using two-way raking to ensure their consistency.

The difference between preliminary and final estimates lies in the timeliness of the sources used to estimate this component. Since the subprovincial estimates of emigrants are adjusted to provincial and territorial estimates, the level of the subprovincial estimates will be the same.

Net temporary emigration

At the subprovincial level, provincial and territorial net temporary emigration estimates by age and sex are broken down based on the subprovincial distribution of emigrants. They are then adjusted to the provincial and territorial totals using two-way raking to ensure their consistency.

The difference between preliminary and final estimates lies in the timeliness of the net temporary emigration estimates.

Returning emigrants

As with immigrants and emigrants, the number of returning emigrants at the subprovincial level is derived from the T1FF. Because the estimates are available only by broad age groups (0-17, 18-24, 25-44, 45-64, 65+), they are broken down by age and sex based on the provincial or territorial distribution. They are then adjusted to the provincial and territorial totals using two-way raking to ensure their consistency.

The difference between preliminary and final estimates lies in the timeliness of the sources used to estimate this component. Since the subprovincial estimates of returning emigrants are adjusted to provincial and territorial estimates, the level of the subprovincial estimates will be the same.

Interprovincial migration

Interprovincial migration by broad age groups and sex for subprovincial regions is derived from the T1FF for each subprovincial region. The estimates by broad age groups and sex are broken down by age based on distributions stemming from the most recent census or NHS (starting in 2011) mobility question on place of residence one year ago. Since 2011/2012, NHS distributions have been modelled to minimize the impact of outliers found in some subprovincial regions, mostly for smaller geographies. Subprovincial estimates are then adjusted to the provincial and territorial totals using two-way raking to ensure their consistency.

The difference between preliminary and final estimates lies in the timeliness of the sources used to estimate this component. Since the subprovincial estimates of interprovincial migrants are adjusted to provincial and territorial estimates, the level of the subprovincial estimates will be the same.

Intraprovincial migration

As with interprovincial migration, the components of intraprovincial migration by broad age groups and sex are derived from the T1FF for each subprovincial region. The estimates by broad age groups and sex are broken down by age based on distributions stemming from the most recent census or NHS (starting in 2011) mobility question on place of residence one year ago. Since 2011/2012, NHS distributions have been modelled to minimize the impact of outliers found in some subprovincial regions, mostly for smaller geographies.

These sources are used for both preliminary and final estimates.

The difference between preliminary and final estimates lies in the availability of the T1FF data used to estimate this component.

Since there are no reliable data sources for preliminary intraprovincial migration estimates, the data for the most recent year, for which final estimates are available, are used. The assumption that intraprovincial migratory behaviours for the current year are similar to those for the previous year for which final estimates are available is adopted.

8.3 Intercensal population estimates for subprovincial regions

Intercensal estimates for census divisions (CDs), census metropolitan areas (CMAs) and economic regions (ERs) are produced much in the same manner as intercensal estimates at the provincial and territorial level (see Chapter 1 for information on the methods). There are three main steps in the production of intercensal estimates:

- the correspondence of the geographic boundaries between the two censuses
- calculation of the error of closure
- linear distribution of the error of closure (residual deviation).

To ensure geographical concordance, the base populations and components of population growth must be adjusted according to geographical boundaries at the time of the most recent census. For areas whose geographical boundaries changed between the two censuses (as measured by the SGC), historical conversion factors are used based on population transfers at the census subdivision level during the most recent intercensal period. In general, corrections to CDs, CMAs and ERs are minor.

Error of closure is defined as the difference between postcensal population estimates on census day and the population enumerated in that census adjusted for census net undercoverage (CNU). The error of closure is spread evenly over the intercensal period, based on the number of days in each month. Intercensal estimates by age and sex are adjusted the same way (i.e., by distributing the error of closure evenly across the age and sex cohorts). As with postcensal estimates, the intercensal subprovincial estimates by age and sex are adjusted to provincial and territorial estimates using two-way raking to ensure their consistency.

8.4 Estimations de la population des subdivisions de recensement

Census subdivision (CSD) is the general term for municipalities (as determined by provincial and territorial legislation) or areas treated as municipal equivalents for statistical purposes (e.g., Indian reserves, Indian settlements and unorganized territories). They are classified into 54 types according to official designations adopted by provincial, territorial or federal authorities. Two exceptions are subdivision of unorganized (SNO) in Newfoundland and Labrador, and subdivision of county municipality (SC) in Nova Scotia, which are geographic areas created as equivalents for municipalities by Statistics Canada, in cooperation with those provinces, for the purpose of disseminating statistical data.

Because there are no components of sufficient quality at that level, the population of CSDs is not estimated using the cohort component method, as is done with census divisions (CDs), census metropolitan areas (CMAs), provinces and territories. The method used consists of applying the CD growth rate to the base population of its CSDs. Two sets of data are necessary in this method: the base population of CSDs and the annual population estimates for CDs.

8.4.1 Base population of census subdivisions

A base population is the population at the beginning of a period, used as a starting point for the estimation process. At the CSD level, the base population is the population count by age and sex for five-year censuses, adjusted for coverage errors.[44] The census data are adjusted as follows:

- adjustment to take population reviews into account. Because there are no population reviews by age and sex, calibration by age and sex is carried out to ensure consistency with the total counts adjusted for population reviews.
- adjustment of the population for census net undercoverage (CNU). Given that coverage studies are not designed to produce subprovincial-level estimates of CNU, provincial and territorial rates by age and sex are used.
- addition of independent population estimates for incompletely enumerated Indian reserves by age and sex.

44. The base population universe is discussed in Chapter 2.

8.4.2 Postcensal population estimates for census subdivisions

CSD populations are estimated in four steps: (1) CD-level population estimates by age and sex are first calculated; (2) CD-level growth rates are calculated by age and sex; (3) these growth rates by age and sex are applied to the corresponding CSD population estimates; (4) CSD-level estimates are adjusted to ensure consistency with the CD-level estimates by age and sex.

Step 1 – Estimating the CD population

The first step consists of estimating the CD-level populations by age and sex using the component method described previously in this chapter.

Step 2 – Calculating CD growth rates by age and sex

The second step involves calculating population growth rates by age and sex for each CD. The formula is as follows:

Equation 8.5: $$\mathrm{GR\_{j}CD}^{a}_{(t,\,t+1)} = \frac{\mathrm{Pop\_{j}CD}^{a}_{(t+1)} - \mathrm{Pop\_{j}CD}^{a}_{(t)}}{\mathrm{Pop\_{j}CD}^{a}_{(t)}}$$

where

$\mathrm{GR\_{j}CD}^{a}_{(t,\,t+1)}$ = population growth rate of CD j at age a for period t and t +1;

$\mathrm{Pop\_{j}CD}^{a}_{(t)}$ = population of CD j at age a and time t.

Step 3 – Estimating postcensal population estimates of CSDs

The third step consists of applying the CD-level growth rates by age and sex to the corresponding CSD base populations. The formula is as follows:

Equation 8.6:

$$\mathrm{Pop}_{\mathrm{\_{j}CD{i}}}\mathrm{CSD}^{a}_{(t+1)} = \mathrm{Pop}_{\mathrm{\_{j}CD{i}}}\mathrm{CSD}^{a}_{(t)} + \left(\mathrm{Pop}_{\mathrm{\_{j}CD{i}}}\mathrm{CSD}^{a}_{(t)} \times \mathrm{GR\_{j}CD}^{a}_{(t,\,t+1)}\right)$$

where

$\mathrm{Pop}_{\mathrm{\_{j}CD{i}}}\mathrm{CSD}^{a}_{(t)}$ = population of CSD i of CD j at age a and time t;

$\mathrm{GR\_{j}CD}^{a}_{(t,\,t+1)}$ = population growth rate of CD j at age a for period t and t+1.

Step 4 – Calibrating the CSD-level population estimates to ensure consistency with the CD population estimates

Finally, to ensure consistency between the CSD and CD population estimates by age and sex, the CSD population estimates are adjusted by age and sex using two-way raking.

Chapter 9

Population estimates by age, sex, marital status and legal marital status

The 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 or legal marital status is available only at the national, provincial and territorial levels. This chapter presents the methods used to produce population estimates disaggregated by age, sex, marital status and legal 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, net non-permanent residents, emigration, returning emigration, net temporary emigration and interprovincial migration) that can be directly linked to persons belonging to the same birth cohort (i.e., persons who were born during the same period or year). Different components require different approaches, based on the nature of the data used to generate the estimates. 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:[45]

- Births and deaths using vital statistics;
- Immigration and non-permanent residents using data from Citizenship and Immigration Canada (CIC);
- Emigration distributed by age and sex using the data by five-year age group, sex, province and territory from T1FF[46] files adjusted for the coverage. We distribute these estimates by single year of age using Sprague coefficients;
- Net temporary emigration distributed by age and sex using emigration distributions;
- Returning emigrants distributed by age and sex using the most recent National Household Survey (NHS) data on mobility data one year ago, after excluding non-permanent residents and immigrants;
- Interprovincial migration by age and sex derived from T1FF family file by Income Statistics Division and counts from the last available NHS (one-year mobility variable).

45. For the methodology used to produce estimates by age and sex for each component of demographic growth, refer to their respective chapters in this report.
46. The T1 family file (T1FF) is derived from the Canada Revenue Agency (CRA) T1 file by Income Statistics Division of Statistics Canada.

9.1.2 Levels of estimates

The difference between preliminary[47] 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.

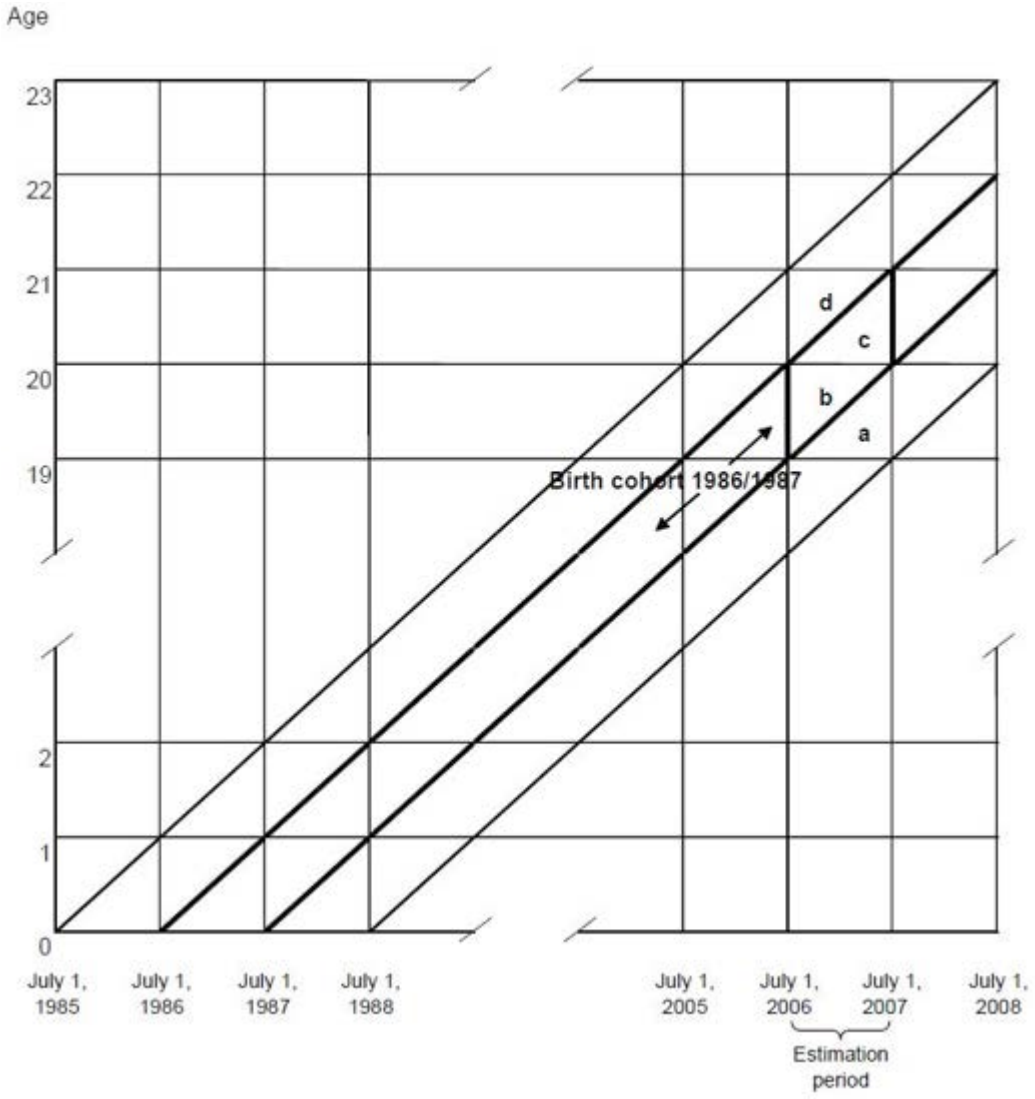
Annual estimates

Estimates of population by age and sex are published annually with July 1 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 aging 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 births, immigration, emigration, net temporary emigration, returning emigration, non-permanent residents and interprovincial migration 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.

47. Unless otherwise noted, the term "preliminary" includes both preliminary and updated estimates when it refers to components.

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, 2006, who belong to the cohort born between July 1, 1986 and June 30, 1987 (inclusive). The demographic events experienced by this cohort during the estimation period are represented by triangles b and c.

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

For each sex, by province and territory:

At age 0:[48]

Equation 9.1:

$$P^{0}_{(t+1)} = B_{(t,t+1)} - D^{-1}_{(t,t+1)} + I^{-1}_{(t,t+1)} - (E^{-1}_{(t,t+1)} + \Delta TE^{-1}_{(t,t+1)}) + RE^{-1}_{(t,t+1)} + \Delta NPR^{-1}_{(t,t+1)} + \Delta N^{-1}_{(t,t+1)}$$

From 1 to 99 years:

48. Age -1 refers to the cohort who had a demographic event at age 0 during the interval (t, t+1) but who was not born at the beginning of the interval.

Equation 9.2:

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

For 100 years and over:

Equation 9.3:

$$P_{(t+1)}^{100+} = P_t^{99+} - D_{(t,t+1)}^{99+} + I_{(t,t+1)}^{99+} - (E_{(t,t+1)}^{99+} + \Delta TE_{(t,t+1)}^{99+}) + RE_{(t,t+1)}^{99+} + \Delta NPR_{(t,t+1)}^{99+} + \Delta N_{(t,t+1)}^{99+}$$

where

$(t, t+1)$ = interval between times t and t+1;

a = age;

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

P_t = base population at time t (census adjusted for CNU[49] or most recent estimate);

B = number of births;

D = number of deaths;

I = number of immigrants;

E = number of emigrants;

ΔTE = net temporary emigration;

RE = number of returning emigrants;

ΔNPR = net non-permanent residents;

ΔN = net interprovincial migration.

Annual population estimates by single year of age and sex for persons aged 0 to 99, and 100 years and over are available from 2001 and onwards at the national, provincial and territorial level. Previously, the upper limit of the age category was 90 years and over.

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 after being adjusted for census net undercoverage (CNU). Like the total population by province or territory, intercensal population by age and sex are adjusted by distributing the error of closure uniformly across the age-sex cohorts. Refer to Chapter 1 for further details.

9.3 Subprovincial postcensal and intercensal estimates by age and sex

Postcensal population estimates by age and sex for census divisions (CDs) and census metropolitan areas (CMAs) 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. A different method called the census division's (CD) aggregate method is used to produce population estimates by

49. Unless otherwise noted, the adjustment for the census net undercoverage (CNU) also includes the incompletely enumerated Indian reserves.

age and sex for economic regions (ERs). Descriptions of the methods used to estimate the populations by age and sex for CMAs, CDs and ERs are provided in Chapter 8. At the subprovincial level, annual population estimates by age and sex are available for ages 0 to 89 and ages 90 and over.

Special methods for preliminary postcensal estimates by age and sex are applied for CDs, CMAs and ERs in 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, marital status and legal 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. However, estimates for the marital status exist since 1971. Estimates of marital statuses are not produced for subprovincial levels.

9.4.1 Definition of different groupings by marital status or legal marital status

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. A person's legal 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 legally 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, separated, widowed or divorced.

Legal marital status	
Single (never legally married)	Includes persons who have never married (including all persons less than 15 years of age). Those who live with a common-law partner are included in this category
Married (and not separated)	Includes persons whose opposite- or same-sex spouse is living, unless the couple is separated or a divorce has been obtained. Also included are persons in civil unions.
Separated	Includes persons currently legally married but who are no longer living with their spouse (for any reason other than illness, work or school) and have not obtained a divorce. Those who live with a common-law partner are included in this category.
Widowed	Includes persons who have lost their legally-marriedspouse through death and have not remarried. Those who live with a common-law partner are included in this category.
Divorced	Includes persons who have obtained a legal divorce and have not remarried. Those who live with a common-law partner are included in this category.

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.

Marital Status	
Single (never legally married)	Includes persons who have never married (including all persons less than 15 years of age). Those who live with a common-law partner are not included in this category.
Married (and not separated)	From 1971 to 1990, the category married (and not separated) includes persons whose opposite- or same-sex spouse is living, unless the couple is separated or a divorce has been obtained. Also included are persons in civil unions and common-law unions. As of 1991, legally married persons are included in this category; persons living with a common-law partner are no longer included in the married category. For the legal marital status, common-law unions are found in each of the categories other than married. For marital status, they constitute a separate category.
Separated	Includes persons currently legally married but who are no longer living with their spouse (for any reason other reasons than illness, work or school) and have not obtained a divorce. Those who live with a common-law partner are not included in this category
Living in common law	Includes persons who are living with a person of the opposite sex or of the same sex as a couple but who are 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.
Widowed	Includes persons who have lost their legally-married spouse through death and who have not remarried. Those who live with a common-law partner are not included in this category.
Divorced	Includes persons who have obtained a legal divorce and have not remarried. Those who live with a common-law partner are not included in this category.

9.5 Postcensal population estimates by marital status and legal marital status

In the past, the Demography Division used the component method to produce population estimates by marital status and legal marital status. Since marriage and divorce data is no longer available, the division had to modify its estimation method. As the overall picture of marital status evolves slowly, the division has opted to use proportions taken from the census. For every census, a series of proportions by age, sex, marital status and legal marital status will be produced. They will be kept constant for the entire postcensal period and will be applied to the annual population estimates by age and sex.

9.5.1 Censal population estimates

The base numbers used to create the proportions come from census counts by age, sex, marital status and legal marital status readjusted for CNU and a particular adjustment made for the population aged 15 to 19 years. For more information regarding the CNU, please refer to Chapter 2.

9.5.1.1 Adjustment for those aged 15 to 19 years

For the past few censuses, we have realised that the census numbers for marital status adjusted for CNU for the 15 to 19 year age group overestimate the number of persons widowed, divorced or married. Comparisons with other data sources, such as vital statistic files (to determine a person's age when they were married or divorced), or Citizenship and Immigration Canada files (to determine the marital status of immigrants or non-permanent residents) do not corroborate the higher counts found in the census. As a result, we decided to include an adjustment for this particular age group.

In order to do so, we use the latest postcensal estimates (July 1st, 2006) that we obtained using the components method in order to correct the adjusted census counts.

We calculate a series of weights for marital status estimates, and another for legal marital status for persons married, separated, divorced and widowed for each age between 15 and 19 years, for both sexes and for each province and territory.

Equation 9.4:

$$\mathrm{Weight}_{ex}^{a} = \frac{P_{ex}^{a}}{\sum_{ex} P_{ex}^{a}}$$

where

a = age going from 15 to 19 years;

ex = marital status or legal marital status for category x. The value of x representing persons either married, separated, divorced or widowed;

P = postcensal population estimates derived from the components method given the marital status or legal marital status.

As we don't have postcensal estimates from the components method for separated persons, we must make an additional adjustment for persons married and separated.

The weight for married persons is calculated as follows:

Equation 9.5:

If $\mathrm{CenA}^{a}_{M+S} = 0;$

$$\mathrm{Weight}^{a}_{M} = \frac{\mathrm{CenA}^{15-19}_{M}}{\mathrm{CenA}^{15-19}_{M+S}} * \frac{\mathrm{P}^{a}_{M+S}}{\sum_{ex} \mathrm{P}^{a}_{ex}}$$

If not $\mathrm{Weight}^{a}_{M} = \frac{\mathrm{CenA}^{a}_{M}}{\mathrm{CenA}^{a}_{M+S}} * \frac{\mathrm{P}^{a}_{M+S}}{\sum_{ex} \mathrm{P}^{a}_{ex}}$

The weight for separated persons is calculated as follows:

Equation 9.6: $\mathrm{Weight}^{a}_{S} = \frac{\mathrm{P}^{a}_{M+S}}{\sum_{ex} \mathrm{P}^{a}_{ex}} - \mathrm{Weight}^{a}_{M}$

After obtaining the weights, we calculate the adjustment for each status.

Equation 9.7:

If $\mathrm{Adjustment}^{a}_{ex} = \mathrm{Weight}^{a}_{ex} * \sum_{ex} \mathrm{Cen}^{a}_{ex} - \mathrm{CenA}^{a}_{ex} > 0$

Set equal to 0

If not $\mathrm{Adjustment}^{a}_{ex} = \mathrm{Weight}^{a}_{ex} * \sum_{ex} \mathrm{Cen}^{a}_{ex} - \mathrm{CenA}^{a}_{ex}$

where

a	=	age going from 15 to 19 years;
ex	=	marital status or legal marital status for category x. The value of x representing persons either married, separated, divorced or widowed;
P	=	postcensal population estimates derived from the components method given the marital status or legal marital status;
M	=	married persons;
S	=	separated persons;
CenA	=	census counts adjusted for CNU.

For people that are single, we sum the adjustment for the other statuses and we change the sign from negative to positive in order to maintain the coherence of the data by age and sex. In other words, the adjustment redistributes people from the other marital statuses and adds them to the count of single persons.

Subsequently, this adjustment is applied to the census counts adjusted for CNU (CenA).

Regardless of the adjustments performed, the coherence between the marital status estimates and those for legal marital status must be assured in order to obtain the censal estimates. The difference between marital status estimates and legal marital status estimates must always be inferior or equal to 0 for statuses other than married. This difference can be explained by the fact that persons living in common law are redistributed from their legal marital status in order to create the common-law category under marital status.

The censal estimates by marital status and legal marital status are used to calculate the intercensal and postcensal estimates.

9.6 Postcensal population estimates by marital status and legal marital status

9.6.1 Estimation methods

In order to estimate the postcensal population estimates by age, sex, marital status and legal marital status we need the censal estimates by age, sex, marital status and legal marital status (adjusted for CNU and a demographic adjustment for those aged 15 to 19 years) and the postcensal estimates by age and sex.

The postcensal estimates by age and sex are distributed by marital status and legal marital status given the distribution of the censal estimates in the following way:

Equation 9.8:
$$P_t^{a,s,ml,m} = \frac{CE_\alpha^{a,s,ml,m}}{\sum_{ml,m} CE_\alpha^{a,s,ml,m}} * PAS_t^{a,s}$$

If the denominator is null for age a, we will use age a-1.

where

$P_t^{a,s,ml,m}$ = postcensal estimates on July 1 of year t, for age a, sex s, legal marital status ml and marital status m;

$CE_\alpha^{a,s,ml,m}$ = censal estimates on date α for age a, sex s, legal marital status ml and marital status m;

$PAS_t^{a,s}$ = postcensal estimate on July 1 of year t, for age a and sex s.

The possible combinations of legal marital status by marital status are:

Legal marital status	Marital status
Single	Single
	Single and living in common law
Married	Married
Separated	Separated
	Separated and living in common law
Widowed	Widowed
	Widowed and living in common law
Divorced	Divorced
	Divorced and living in common law

We can derive two series of population estimates, one by marital status and the other by legal marital status. People living in common law are derived by finding the sum of people living in common law in each of the categories. As is done with censal population estimates, we must be sure to maintain coherence between the two series.

9.6.2 Levels of estimate

As was previously mentioned, the components method is no longer used to produce population estimates by marital status. Rather, these estimates are based on the marital status proportions taken from the Census and the population estimates by age and sex, which in turn were produced using the components method. Therefore, the difference between the preliminary and final postcensal population estimates by marital status or legal marital status can be found in the revision level of the components that were used to estimate the population by age and sex. If all the components are preliminary, we obtain preliminary postcensal estimates (PP). If they are all final components, the estimates become final postcensal (PD). For any other combination, we have updated postcensal estimates (PR).

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, marital status and legal marital status from the two most recent censuses are used to derive intercensal estimates of population by marital status and 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, marital status and legal marital status.

The calculations are as follows:

Equation 9.9:

$$\mathrm{IP}_{t}^{a,s,ml,m} = \left[\frac{\mathrm{CE}_{\alpha}^{a,s,ml,m}}{\sum_{ml,m} \mathrm{CE}_{\alpha}^{a,s,ml,m}} + \left(\frac{t-\alpha}{\beta-\alpha} \right) \times \left(\frac{\mathrm{CE}_{\beta}^{a,s,ml,m}}{\sum_{ml,m} \mathrm{CE}_{\beta}^{a,s,ml,m}} - \frac{\mathrm{CE}_{\alpha}^{a,s,ml,m}}{\sum_{ml,m} \mathrm{CE}_{\alpha}^{a,s,ml,m}} \right) \right] \times \mathrm{IP}_{t}^{a,s}$$

where

$\mathrm{IP}_{t}^{a,s,ml,m}$ = intercensal population estimates at date t for age a, sex s, legal marital status ml and marital status m;

$\mathrm{CE}_{\alpha}^{a,s,ml,m}$ = censal population estimates at date α for age a, sex s, legal marital status ml and marital status m;

$\mathrm{CE}_{\beta}^{a,s,ml,m}$ = censal population estimates at date β for age a, sex s, legal marital status ml and marital status m;

α = date for the first census (census at the beginning of the period);

β = date for the second census (census at the end of the period);

t = date of the intercensal estimate on July 1 of year t;

$\mathrm{IP}_{t}^{a,s}$ = intercensal population estimate on July 1 of year t, for age a and sex s.

We can derive two series of population estimates, one by marital status and the other by legal marital status. People living in common law are derived by finding the sum of people living in common law in each of the categories. As is done with postcensal population estimates, we must be sure to maintain coherence between the two series.

Chapter 10

Estimates of census families, economic entities and households

In addition to population estimates, Statistics Canada produces annual estimates of census families for Canada, provinces and territories. These estimates are available from 1986 onwards. Estimates of the number of census families can be either postcensal or intercensal. With the exception of the territories, estimates of census families are available by family type, sex and age group of parents, family size and children's age group.

The Demography Division has restructured its program in order to estimate the census families. This is different from the components method[50] used in previous years. The estimation method described in this chapter is the headship rates method. The series produced by this method goes back to 2006. However, readers can refer to the 2007 version of this report for further details on the components method.

The Demographic Estimates Program also produces estimates of economic entities and private households at the national, provincial and territorial levels by the headship rates method. Estimates for these two analytical units are used internally[51] and are available on request. In this chapter, reference to the development of private households and economic entities are also presented as a harmonized estimation of these three demographic indicators can be attained with the new headship rates method.

10.1 Definitions and relevant concepts

For the purpose of generating these estimates, the definitions of the three concepts are the same as those in the 2011 Census.[52]

A 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. A couple may be of opposite or same sex.[53] Children in a census family include grandchildren living with their grandparent(s) but with no parents present. For the estimates, married couples and couples living common-law are grouped into a single category.

An economic entity is derived from the concept of economic family. It includes economic families and persons not living in an economic family (who accordingly constitute economic entities of size 1). According to the census, an economic family is defined as a group of two or more persons who live in the same dwelling and are related to each other by blood, marriage, common-law, adoption or a foster relationship. A couple may be of opposite or same sex.

A household refers to a person or a group of persons (other than foreign residents) who occupy the same dwelling and do not have a usual place of residence elsewhere in Canada. It may consist of a family group (census family) with or without other persons, of two or more families sharing a dwelling, of a group of unrelated persons, or of one person living alone. Household members who are temporarily absent on Census Day (e.g., residing temporarily elsewhere) are considered as part of their usual household. For census purposes, every person is a member of one and only one household. Unless otherwise specified, all data in household reports are for private households only.

10.2 Data sources

Four sources of statistical data, all described previously in this document, are needed to produce estimates of these three entities: tax data from the T1 Family File (T1FF), census data, data on census net undercoverage from coverage studies, and population estimates. Table 10.1 shows the sources, references and role of the data used for estimating these analytical units.[54]

50. For details on the components method, refer to Statistics Canada 2007, Population and Family Estimation Methods at Statistics Canada, Catalogue number 91-528-XIE, Statistics Canada.
51. These estimates are used in the calculation of demographic, social and economic indicators. They are used for planning, program evaluation and base population for various surveys and studies. In addition, these data are also used in calculating of weights for use in Statistics Canada surveys.
52. For more details, see the 2011 Census Dictionary
53. Before the implementation of the headship rates method in 2011, same-sex couples were excluded in the estimation of census families, since it was impossible to establish trends with regard to the increase of this group by the components method.
54. Analytical units refer to census families, economic entities and households.

Table 10.1
Sources, references and role of data used to produce estimates of census families, economic entities and households

Source	Reference	Role
T1 Family File	Income Statistics Division, Statistics Canada	Creation of census correction coefficients
Census	Statistics Canada	Creation of census correction coefficients Distribution of size of analytical units of 6 or more persons
Census coverage studies	Social Survey Methods Division, Statistics Canada	Adjustment of census net undercoverage
Population estimates	Demography Division, Statistics Canada	Adjustment of census net undercoverage

10.3 Postcensal estimates of census families, economic entities and households

Annual estimates of census families, economic entities and private households are based on T1FF tax data with different adjustments. Even though the T1FF data does not create economic entities and private households, it can be used to estimate these analytical units. The method can be divided into three steps: correct the biases inherent in the T1FF, adjust the reference date and adjust the census coverage.

Step 1 – Correction of biases inherent in the T1FF

The coverage provided by T1FF tax data, while excellent, is not complete. Also, some operational definitions in the T1FF do not directly correspond to census definitions. Therefore, it is necessary to make corrections to the T1FF by comparing the number of heads of census families in the T1FF with the number in the census for each of the three analytical units. The way of determining heads varies depending on the concept:

- Census families:
    - Couples: each spouse shares the role of family head and is therefore equivalent to half a head;
    - Single-parent families: the parent is the head;
    - Non-family: each individual is a head.
- Economic entities:
    - Couples: each spouse is equivalent to half a head, as in the case of census families;
    - Single-parent families: the parent is the head;
    - Non-family: each individual is a head.
- Private Households:
    - The primary household maintainer is the head (a household may have more than one maintainer).

It is therefore possible to create a census correction coefficient for each combination of characteristics, namely province or territory of residence, age group, sex, type of census family, age group of spouse, sex of spouse, size of census family and age group of children. The way to calculate this coefficient depends on the situation, as Table 10.2 shows.

Table 10.2
Possible situations for calculating census correction coefficient

Situation	$H_{CF}^{T1FF}(p, a, s, fam, a', s', sf, a'')$	$H_i^{Cen}(X)$	$E_i(X)$	Level
1	>0	>0	$\frac{H_i^{Cen}(X)}{H_{CF}^{T1FF}(p,a,s,fam,a',s',sf,a'')}$	8
2	=0	=0	0	0
3	>0	=0	0	0
4	=0	>0	See below	

where

$E_i(X)$ = Census correction coefficient is the ratio of census heads of census families, economic entities and households by characteristics (X) to T1FF heads of census families by (p,a,s,fam,a',s',sf,a'');

H^{T1FF} = T1FF heads;

H^{Cen} = census heads;

X = p, a, s, fam, a', s', sf, a'', efam, se, sh;

p = province or territory of residence;

a = age group;

s = sex;

fam = type of census family;

a' = age group of spouse;

s' = sex of spouse;

sf = size of census family;

sh = size of household;

a'' = age group of children;

ee = type of economic entity;

se = size of economic entity;

CF = census families;

EE = economic entities;

HSD = households.

i = CF, EE or HSD;

In Table 10.2, 'Level' refers to census family characteristics. For example, in the first situation, the correction coefficient is available for each combination of the eight characteristics. In the second and third situations, a correction coefficient of 0 indicates no estimates of heads in the census.

In the fourth situation, the census correction coefficient is calculated using a special method involving the use of higher 'level' T1FF counts. The situation represents a very small portion of the estimates.

The correction coefficient is also used to estimate economic entities and private households from T1FF. These census correction coefficients are calculated each census year and are assumed to remain constant for the entire postcensal period. They represent both an adjustment of coverage and an adjustment of conceptual bias which is illustrated in the following decomposition:

Equation 10.1: $$E_i(X) = \text{bias} \times \text{relationship} = \frac{H_i^{Cen}}{H_i^{T1FF}} \times \frac{H_i^{T1FF}}{H_{CF}^{T1FF}} = \frac{H_i^{Cen}}{H_{CF}^{T1FF}}$$

Each year, it will be possible to correct the biases inherent in the T1FF simply by multiplying the number of heads in the T1FF by the census correction coefficients, taking the level of each coefficient into account.

Equation 10.2: $$CH_i(X) = E_i(X) \times H_{CF}^{T1FF}(X)$$

where

$CH_i(X)$ = corrected number of heads in the T1FF.

Step 2 – Adjustment of the reference date

The T1FF is based on tax data, mostly collected no later than the end of April of the following year and often refer to the situation on December 31, whereas the demographic estimates generally have different reference dates (such as July 1). The census correction coefficient in Step 1 adjusts the reference date from December 31 back to Census Day. Here, the correction applied to the T1FF data set adjusts the reference date from Census Day of the current year to July 1. This time lag is factored in using headship rates, constructed with population estimates. Since census net undercoverage will be taken into account in the next step, the population estimates used for the headship rates exclude the adjustment for census net undercoverage. For an estimate at July 1, the adjustment of the reference date is as follows:

Equation 10.3: $$CH_i^{Jul1}(X) = \frac{CH_i^{CD}(X)}{Unadjust\_pop_{CD}^{Dem}(p,a,s)} \times Unadjust\_pop_{Jul1}^{Dem}(p,a,s)$$

where

$CH_i^{Jul1}(X)$ = corrected number of heads in the T1FF at July 1;

$CH_i^{CD}(X)$ = corrected number of heads in the T1FF at Census day;

$Unadjust\_pop_{CD}^{Dem}(p,a,s)$ = $Pop_{CD}^{Dem}(p,a,s) - Pop\_CNU(p,a,s)$;

$Unadjust\_pop_{Jul1}^{Dem}(p,a,s)$ = $Pop_{Jul1}^{Dem}(p,a,s) - Pop\_CNU(p,a,s)$;

Pop_CNU = net undercoverage of census population;

CD = Census day of the current year.

Step 3 – Adjustment of census coverage

Since some individuals are not counted during the census, a certain number of census families, economic entities and households is also omitted. A correction is therefore needed to take these omitted analytical units into account.

Social Survey Methods Division (SSMD) provided estimates of 2011 census net undercoverage for the households. Demography Division used these estimates to also calculate the census net undercoverage for census families and economic entities.

Two assumptions are being made in order to use this method of calculation. First assumption is that the rate of net undercoverage for census families, economic entities and households is the same. The second is that the characteristics of the 2011 census net undercoverage are the same as 2006.

Final equation

By combining all three steps above, it becomes possible to construct the final equation for producing estimates. For an estimate at July 1, the equation is:

Equation 10.4:

$$\mathrm{EST_i(X)} = \frac{\mathrm{H_{CF}^{T1FF}(X) \times E_i(X)}}{\mathrm{Unadjust\_pop_{CD}^{Dem}(p,a,s)}} \times \mathrm{Unadjust\_pop_{Jul1}^{Dem}(p,a,s)} + \mathrm{EST\_CNU_i(X)}$$

where

$\mathrm{EST_i(X)}$ = estimated number of census families, economic entities or households;

$\mathrm{EST\_CNU_i(X)}$ = estimate of census net undercoverage of census families, economic entities or households.

10.4 Characteristics of census families, economic entities and households

As seen in the previous section, estimates are produced for several characteristics and are estimated at the same time as the number of analytical units. These characteristics are:

- province or territory of residence;
- age group of heads;
- sex of head(s);
- type of census family (couple, single-parent family, non-census family);
- size of census family (2, 3, 4, 5, 6 and more members);
- age group of children;
- type of economic entity (couple, single-parent family, non-economic family);
- size of economic entity (2, 3, 4, 5, 6 and more members);
- size of household (2, 3, 4, 5, 6 and more members).

10.5 Postcensal estimates of number of persons living in analytical units

In addition to producing estimates of the number of census families, economic entities and households, it is also possible to produce an estimate of the number of persons associated with each of these analytical units. This estimate is produced in two parts: first, for units of which the size is 1 to 5 persons and second, for units of size 6 or more persons.

Size of 1 to 5 persons

For these units, the estimate is direct:

Equation 10.5: $$\mathrm{IND_i(X|Size \in [1,5])} = \mathrm{EST_i(X|Size \in [1,5])} \times \mathrm{Size_i(X)}$$

where

$\mathrm{IND_i(X|Size \in [1,5])}$ = number of people in each analytical unit whose size is 1 to 5 people;

$\mathrm{EST_i(X|Size \in [1,5])}$ = number of analytical units;

$\mathrm{Size_i(X)}$ = size of the analytical unit.

Size of 6 or more persons

For these units, since their actual size is suppressed in the construction of the model, we use provincial or territorial distributions from the last census to estimate the average size for analytical units of 6 or more persons, and then apply it to the estimates of each analytical unit. The number of persons in each analytical unit with size of 6 or more persons can be obtained by using the following equation:

Equation 10.6: $$\mathrm{IND_i(X|Size = 6+)} = \mathrm{EST_i(X|Size = 6+)} \times \mathrm{AverageSize_i(p)}$$

where

$\mathrm{IND_i(X|Size = 6+)}$ = number of people in each analytical unit whose size is 6 or more persons;

$\mathrm{EST_i(X|Size = 6+)}$ = estimate of each analytical unit;

$\mathrm{AverageSize_i(p)}$ = average size of the analytical units with 6 or more persons.

Estimated number of persons

We can obtain the total number of persons living in each analytical unit by summing the two results obtained.

Equation 10.7: $$\mathrm{IND_i(X) = IND_i(X|Size} \in [1,5]) + \mathrm{IND_i(X|Size = 6+)}$$

10.6 Level of estimates

Postcensal estimates may be preliminary[55] or final; the difference being in the level of timeliness of source files.

- Preliminary postcensal estimates:
  - Preliminary postcensal population estimates;
  - Projection of T1FF for the fiscal year corresponding to the reference year (for example, the projection of T1FF for the 2014 fiscal year for estimates at July 1, 2014);
- Final postcensal estimates:
  - Final postcensal population estimates;
  - T1FF for the fiscal year corresponding to the reference year (for example, the T1FF for the 2014 fiscal year for estimates at July 1, 2014).

10.7 Intercensal estimates of census families

Once the postcensal cycle is completed, intercensal estimates can be calculated. The method remains the same as for postcensal estimates; only the parameters of the model are changed. The census correction coefficients, the average size for the units consisting of 6 or more persons and the units added to adjust for census net undercoverage are interpolated linearly, while the population estimates used are intercensal.

55. Unless otherwise noted, the term "preliminary" includes both preliminary and updated estimates when it refers to components.

Glossary

Age

Age as of July 1.

Census coverage

Census net undercoverage (CNU): Difference between undercoverage and overcoverage.

Overcoverage: Number of persons or families who were counted more than once in the census.

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

Census division (CD)

Census division (CD) is the general term for provincially legislated areas (such as county, *municipalité régionale de comté* and regional district) or equivalent regions. Census divisions are intermediate geographic areas between the municipalities (census subdivision) and the province or the territory.

In Newfoundland and Labrador, Manitoba, Saskatchewan, Alberta, Yukon, Northwest Territories and Nunavut, provincial and territorial laws do not provide for these administrative geographic areas. Therefore, Statistics Canada in cooperation with these provinces and territories, has created equivalent areas called census divisions for the purpose of disseminating statistical data. In Yukon, the census division is equivalent to the entire territory.

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. A couple may be of opposite or same sex. The children in a census family include the grandchildren who live in the household of at least one of their grandparents, in the absence of the parents.

Census metropolitan area (CMA)

A census metropolitan area (CMA) is formed by one or more adjacent municipalities centered on a population centre (known as the core). A CMA must have a total population of at least 100,000 of which 50,000 or more must live in the core. To be included in the CMA, other adjacent municipalities must have a high degree of integration with the core, as measured by commuting flows derived from census place of work data.

Once an area becomes a CMA, it is retained as a CMA even if its total population declines below 100,000 or its core falls below 50,000. Small population centers with a population count of less than 10,000 are called fringe. All areas inside the CMA that are not population centers are rural areas.

All CMAs are subdivided into census tracts.

The CMA of Ottawa-Gatineau (Ontario-Quebec) crosses provincial boundaries. When the geographic level selected is all of Canada, the totals include the CMA on both sides of the provincial border. If a province has been selected, only the part of the CMA in the province chosen is included in the totals.

Census subdivision (CSD)

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

Cohort

Represents a group of persons who have experienced a specific demographic event during a given year. In the case of births, persons born within a specified year are referred to as a generation.

Cohort component approach

This is the 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 organized according to age and sex cohorts. Data required for this method include 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 growth

Any of the classes of events generating population movement variations. Births, deaths and migrations are the components responsible for the variations since they alter either the total population or the age and sex distribution of the population.

Economic region (ER)

An economic region (ER) is a grouping 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 and territories, economic regions are created by agreement between Statistics Canada and the province or 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.

Emigrant

Canadian citizen or immigrant who has left Canada to establish a residence in another country, involving a change in usual place of residence. Emigration may be either temporary or permanent. Where the term is used alone in this manual, it references to a person's permanent emigration which involves severing residential ties with Canada and acquiring permanent residency in another country.

Error of closure

Difference between the postcensal estimate for the same date as the census and the results of the census adjusted for census net undercoverage (including adjustment for incompletely enumerated Indian reserves).

Household

Refers to a person or a group of persons who occupy the same dwelling and do not have a usual place of residence elsewhere in Canada. It may consist of a family group (census family) with or without other persons, of two or more families sharing a dwelling, of a group of unrelated persons, or of one person living alone.

Immigrant

Within the framework of this publication, the terms immigrant, landed immigrant and permanent resident are equivalent. An immigrant refers to a person who is or has ever been a landed immigrant (permanent resident) and who has been granted the right to live in Canada permanently by immigration authorities. Immigrants are either Canadian citizens by naturalization (the citizenship process) or permanent residents under Canadian legislation. Some immigrants have resided in Canada for a number of years, while others have arrived recently. Most immigrants are born outside Canada, but a small number are born in Canada. Also, children born in other countries to parents who are Canadian citizens that reside temporarily in another country are not included in the category as they become Canadian citizens at birth.

Internal migration

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

International migration

International migration represents movement of population between Canada and a foreign country which involves a change of 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 all movements from one province or territory to another involving a change in usual place 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.

Intraprovincial migration or subprovincial migration

Intraprovincial migration or subprovincial migration represents all movements from one region to another within the same province or territory involving a change in usual place 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.

Legal marital status

Indicates the legal conjugal arrangement of a person. Estimates are presented in the following categories: single (never legally married), married (and not separated), separated, divorced, widowed. Common law unions do not constitute a distinct category of the legal marital status. Persons living in common law unions are distributed among all the legal categories, except for the category: married and not separated.

Single (never legally married): Includes those who have never married (as well as those aged 15 and under). Persons who live in a common law union are included in this category.

Married (and not separated): From 1971 to 1990, the "Married (and not separated)" includes those whose partner of the opposite sex or the same sex is alive, unless the couple has separated or obtained a divorce. Are also included the persons in a civil union, and the persons who are living in a common law union.

As of 1991, only the persons married by law are included in this category. Persons living in common law unions are found in the different categories of the legal marital status, except for the "married (and not separated)" category.

Separated: A married person who no longer lives with his or her partner (for other reasons than illness, work or studies) and who has not obtained a divorce.

Divorced: Person who was obtained a legal divorce and who has not remarried.

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

Marital status

Indicates the conjugal arrangement of a person. Estimates are presented in the following categories: single, married (and not separated), separated, living common law, widowed or divorced.

Single: Includes persons who have never been married (and persons aged less than 15 years old). Persons who live in a common law relationship are not included in this category.

Married (and not separated): For Demography division's estimates, from 1971 to 1990, this category includes those whose partner of the opposite or the same sex is alive, unless the couple has separated or obtained a divorce. Persons with a civil union are also included. Persons who are living with a common law partner are not included in this category.

Separated: Includes those who are married but no longer live with their partner (for reasons other than illness, work or studies) and who have not obtained a divorce. Persons who are living in a common law union are not included in this category.

Common law: Includes persons who are part of a couple with a person of the opposite sex or the same sex but who are not legally married to this person. It also includes situations where the members of the couple are temporarily separated because of illness, work or studies.

Widowed: Persons who have lost their spouse through death and who have not remarried. Persons living in a common law union are not included in this category.

Divorced: Persons who have obtained a legal divorce and who have not remarried. Persons living in a common law union are not included in this category.

Natural increase

Variation of the 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 intraprovincial region. A region can be defined as a census division (CD), an economic region (ER) or a census metropolitan area (CMA).

Net non-permanent residents

Net non-permanent residents represents the variation in the number of non-permanent residents between two dates.

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.

Non-permanent resident

A non-permanent resident is a person who is lawfully in Canada on a temporary basis under the authority of a valid document (work permit, study permit, Minister's permit or refugee) issued for that person along with members of his family living with them. This group also includes individuals who seek refugee status upon or after their arrival in Canada and remain in the country pending the outcome of processes relative to their claim. Note that Citizenship and Immigration Canada uses the term temporary resident rather than non-permanent resident.

Permanent resident

See Immigrant.

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

Censal: Enumerated population adjusted for the net census undercount and partially enumerated Indian reserves. The censal estimate by age and sex can also include a demographic adjustment.

Postcensal: Population estimate produced by using data from the most recent available census adjusted for census net undercoverage (including adjustment for incompletely enumerated Indian reserves) and estimate of the components of demographic 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 census net undercoverage (including adjustment for incompletely enumerated Indian reserves) 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, subtract residual deviation. It can be positive or negative.

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, we calculate births, deaths and immigration rates.

Demographic growth rate 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.

Census net undercoverage rate (CNU): Difference between undercoverage rate and overcoverage rate.

Census overcoverage rate: The ratio of the numbers of persons or families 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.

Census undercoverage rate: The ratio of the estimated number of persons or families not enumerated in the census (who 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 growths 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 (by using the number of days) over the five-year period concerned.

Returning emigrant

Canadian citizen or immigrant having previously emigrated from Canada and subsequently returned to the country to establish a permanent residence.

Reverse Record Check (RRC)

The Reverse Record Check is one of three 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.

Temporary resident

See Non-permanent resident.

Total net migration

Sum of net international and net internal migration.

Vital statistics

Vital Statistics includes all the demographic events (that is to say births, deaths, marriages and divorces) for which there are a legal requirement to inform 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.

References

ANDREEV, Kirill F. 2004. “A Method for Estimating Size of Population Aged 90 and over with Application to the 2000 U.S. Census Data”. *Demographic Research*, Vol. 11, no. 9, p. 235 to 262, http://www.demographic-research.org/volumes/vol11/9/11-9.pdf

BOURBEAU, Robert and André LEBEL. 2000. “Mortality statistics for the oldest-old: an evaluation of Canadian data”. *Demographic Research*, Vol. 2, no. 2, p. 1 to 36, http://www.demographic-research.org/volumes/vol2/2/2-2.pdf

COALE, Ansley J. and Graziella CASELLI. 1990. “Estimation of the Number of Persons at Advanced Ages from the Number of Deaths at Each Age in the Given Year and Adjacent Years”. *Genus*, Vol. 66, no. 1-2, p. 1 to 23, http://www.jstor.org/discover/10.2307/29789011?sid=21105914698401&uid=2&uid=2129&uid=70&uid=35291&uid=3&uid=62&uid=67&uid=407536&uid=3737720&uid=3739448

COALE, Ansley J. and Ellen E. KISKER. 1990. “Defect in Data on Old-Age Mortality in the United-States: New Procedure for calculating Mortality Schedules and Life Tables at the Highest Ages”. *Asian and Pacific Population Forum*, Vol. 4, no. 1, p. 1 to 31.

Deming, W.E., 1943. *Statistical adjustment of data*. Wiley, NY (Dove Publications edition, 1985), 261 p. www.maa.org/press/maa-reviews/statistical-adjustment–of-data

Denton, F. 1971, “Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization”, *Journal of the American Statistical Association*, 66, p.99 to 102.

DICK, Peter. 1995. «Modélisation du sous-dénombrement net dans le recensement du Canada de 1991», Statistics Canada Catalogue number l 12-001-X, Ottawa, Ontario, p. 45 to 54, *Techniques d’enquête*, Vol. 21, no. 1.

DINI, Ercilia and Shayla GOLDRING. 2008. “Estimating the changing population of the ‘oldest old’ “, *Population Trends*, Vol. 132, Summer 2008, Office for National Statistics, U.K., p. 8 to 16.

FAY, Robert E. and Roger A. HERRIOT. 1979. “Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data”, *Journal of the American Statistical Association*, Vol. 81, p. 269 to 277, http://www.jstor.org/stable/2286322?seq=1#page_scan_tab_contents

FELLIGI, Ivan P. and Alan B. SUNTER. 1969. “A Theory of Record Linkage”, *Journal of the American Statistical Association*, Vol. 64, p. 1183 to 1210, http://www.psi.toronto.edu/~yoseph/outbox/2286061.pdf

GAMBINO, Jennifer M. and Peter DICK. 2000. “Small Area Estimation Practices at Statistics Canada”, *Statistics in Transition: Journal of the Polish Statistical Association*, Vol. 4, p. 597 to 610.

HUMAN MORTALITY DATABASE. 2013. *Human Mortality Database* (electronic resource). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany), http:/www.mortality.org.

HUMES, Karen and Victoria A. VELKOFF. 2007. “Centenarians in the United States: information from census 2000”, Document presented in 2007 at the annual meeting of the Population Association of America, New York, NY. http://paa2007.princeton.edu/papers/7066

KANNISTO, Vaino. 1999. “The centenarians”, *The advancing frontier of survival*. Published by the Institut Max Planck pour la recherche démographique, electronic edition, 145 p., http://demogr.mpg.de/Papers/Books/Monograph3/centenarians.htm

KANNISTO, Vaino, Jens LAURITSEN, A. Roger THATCHER and James W. VAUPEL. 1994. “Reductions in Mortality at Advanced Ages: Several Decades of Evidence from 27 Countries”. *Population and Development Review*, Vol. 20, no. 4, p. 793 to 810, http://user.demogr.mpg.de/jwv/pdf/PDR_20.pdf

KESTENBAUM, Bert and B. Renée FERGUSON. 2005. “Number of centenarians in the United States 01/01/1990, 01/01/2000, and 01/01/2010 based on improve Medicare data”. Paper presented at the Population Association of America, Philadelphia, Pennsylvania, March 31 – April 2 2005. http://paa2005.princeton.edu/papers/50718

KRACH, Constance A. and Victoria A. VELKOFF. 1999. “Centenarians in the United States.” U.S. Bureau of Census, *Current Population Reports*, p. 23 to 199. https://www.census.gov/prod/99pubs/p23-199.pdf

LEBEL, André. 1999. *Évolution de la mortalité canadienne aux âges avancés de 1951 à 1995*. Montréal, Département de démographie, Université de Montréal, M.Sc. Thesis.

MICHALOWSKI, Margaret. 1999. *Les Canadiens résidant temporairement à l’étranger : nombre, caractéristiques et méthodes d’estimation*, working document. Statistics Canada. Ottawa, Ontario.

OFFICE FOR NATIONAL STATISTICS. 2013. *Estimates of the Very Old (including Centenarians), 2002-2011, England and Wales*, Statistical Bulletin, Newport, UK, 15 p., http://www.ons.gov.uk/ons/dcp171778_303741.pdf

RAMSAY, T. 2000. *Small Domain Estimation Based on Nonparametric Regression*, working document. Statistics Canada, Ottawa, Ontario.

SIEGEL, Jacob S. and Jeffrey S. PASSEL. 1976. “New Estimates of the Number of Centenarians in the United States”. *Journal of the American Statistical Association*, Vol. 71 Number 355, p. 559 to 566.

SHRYOCK, Henry S., Jacob S. SIEGEL and all. 1976. *Studies in Population: The Methods and Materials of Demography*. Washington: Academic Press, Inc.

SPENCER, G. 1987. “Improvements in the Quality of Census Age Statistics for the Elderly.” *Data for an Aging Population*. Proceedings of the 1987 Public Health Conference on Records and Statistics, July 13–17, Washington, D.C. DHHS Pub. No. (PHS) 88–1214, National Center for Health Statistics, Hyattsville, MD, p. 231 to 235.

STATISTICS CANADA. 2015a. *Quarterly Demographic Estimates*. January to March 2015 Catalogue number 91-002-X, Statistics Canada. Quarterly. http://www.statcan.gc.ca/pub/91-002-x/91-002-x2015001-eng.htm.

- 2015b *Annual Demographic Estimates: Subprovincial Areas, 2014*, Catalogue number 91-214-X, Statistics Canada, Annual. http://www.statcan.gc.ca/pub/91-214-x/91-214-x2015000-eng.htm
- 2015c. *Population Projections for Canada (2013 to 2063), Provinces and Territories (2013 to 2038)*, Catalogue number 91-520-X, Statistics Canada, http://www.statcan.gc.ca/pub/91-520-x/91-520-x 2014001-eng.htm
- 2015d. *2011 Census Technical Report: Coverage,* Catalogue number 98-303-X, Statistics Canada, Ottawa, http://www12.statcan.gc.ca/census-recensement/2011/ref/guides/98-303-x/index-eng.cfm
- 2014. *Annual Demographic Estimates: Canada, Provinces and Territories, 2014*, Catalogue number 91-215-X, Statistics Canada, http://www.statcan.gc.ca/pub/91-215-x/91-215-x2014000-eng.htm
- 2013a. *The Daily*, “Final estimates of 2011 Census coverage”, Catalogue number 11-001E, Statistics Canada, 26 September 2013, http://www.statcan.gc.ca/daily-quotidien/130926/dq130926b-eng.htm
- 2013b . *NHS Profile, 2011*, Catalogue number 98-316, Statistics Canada, http://www12.statcan.ca/nhs-enm/2011/dp-pd/prof/index.cfm?Lang=E
- 2012a. *2011 Census Dictionary*, Catalogue number 98-301-XWE, Statistics Canada, http://www12.statcan.gc.ca/census-recensement/2011/ref/dict/index-eng.cfm
- 2012b. *Population and Dwelling Count Highlight Tables, 2011 Census*, Catalogue number 98-310-X, Statistics Canada, http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/Index-eng.cfm?Lang=Eng

- 2010. *Migration Estimates. User Guide,* Demography Division, Statistics Canada, Ottawa. Internal document.
- 2008. *Mobility and Migration Reference Guide, 2006 Census*, Catalogue number 97-556-GWE2006003, Statistics Canada, Ottawa, http://www.statcan.gc.ca/pub/12-591-x/2006001/step2-etape2/examples-exemples/4113823-eng.htm
- 2007. *Population and Family Estimation Methods at Statistics Canada*. Catalogue number 91-528-XIE, Statistics Canada, Ottawa, 5 p., http://www.statcan.gc.ca/pub/91-528-x/91-528-x2007001-eng.pdf
- 2003. *Population and Family Estimation Methods at Statistics Canada*, Catalogue number 91-528-XIE, Statistics Canada, Ottawa, 168 p., http://www.statcan.gc.ca/pub/91-528-x/91-528-x2003001-eng.pdf

THATCHER, Roger, Vaino KANNISTO and Kirill ANDREEV. 2002. “The survivor ratio method for estimating numbers at high ages”. *Demographic Research*, Vol. 6 Number 1, p. 1 to 18. http://www.demographic-research.org/volumes/vol6/1/6-1.pdf

THATCHER, Roger. 1999. “The demography of centenarians in England and Wales.” *Population Trends* 96, Summer 1999, Office for National Statistics, Newport, U.K., p. 5 to 12.

- 1992. “Trends in Number and Mortality at High Ages in England and Wales.” *Population Studies*, Vol. 46 Number 3, p. 411 to 426, http://dx.doi.org/10.1080/0032472031000146426

US CENSUS BUREAU. 2012. *Centenarian: 2010*. 2010 Census Special Report, p. 14.

VINCENT, Paul. 1951. «La mortalité des vieillards.» *Population*, Vol. 6 Number 2, p. 181 to 204. http://www.jstor.org/stable/1524149?seq=1#page_scan_tab_contents