Statistics Canada Statistique Canada

Catalogue 91-528E

POPULATION ESTIMATION METHODS, CANADA

Demography Division

Canada

Data in Many Forms...

Statistics Canada disseminates data in a variety of forms. In addition to publications, both standard and special tabulations are offered on computer print-outs, microfiche and microfilm, and magnetic tapes. Maps and other geographic reference materials are available for some types of data. Direct access to aggregated information is possible through CANSIM, Statistics Canada's machine-readable data base and retrieval system.

How to Obtain More Information

Inquiries about this publication and related statistics or services should be directed to:

Population Estimates Section,
Demography Division,

Statistics Canada, Ottawa, K1A 0T6 (Telephone: 991-2320) or to the Statistics Canada reference centre in:

St. John's	(772-4073)	Sturgeon Falls	(753-4888)
Halifax	(426-5331)	Winnipeg	(949-4020)
Montréal	(283-5725)	Regina	(780-5405)
Ottawa	(990-8116)	Edmonton	(420-3027)
Toronto	(973-6586)	Vancouver	(666-3691)

Toll-free access is provided in all provinces and territories, **for users who reside outside the local dialing area** of any of the regional reference centres.

Newfoundland and Labrador
1-800-563-4255

Nova Scotia, New Brunswick
and Prince Edward Island
1-800-565-7192

Quebec 1-800-361-2831

Ontario 1-800-268-1151

Manitoba 1-800-282-8006

Saskatchewan 1-800-667-7164

Alberta 1-800-222-6400

British Columbia (South
and Central) 1-800-663-1551

Yukon and Northern B.C.
(area served by
NorthwesTel Inc.) Zenith 0-8913

Northwest Territories
(area served by
NorthwesTel Inc.)
Call collect 403-420-2011

How to Order Publications

This and other Statistics Canada publications may be purchased from local authorized agents and other community bookstores, through the local Statistics Canada offices, or by mail order to Publication Sales, Statistics Canada, Ottawa, K1A 0T6.
1(613)993-7276

Toronto
Credit card only (973-8018)

POPULATION ESTIMATION METHODS, CANADA

Statistics Canada

Demography Division

Published under the
authority of the
Minister of Supply and
Services Canada

March 1987
8-1300-540

Price: Canada, $25.00
Other Countries, $26.50

Payment to be made in
Canadian funds or equivalent

Catalogue 91-528E

ISBN 0-660-12028-3

Ottawa

Version française de cette publication
disponible sur demande (n° 91-528F
au catalogue)

ERRATA

Catalogue 91-528E

POPULATION ESTIMATION METHODS, CANADA

Page 3

The $13.8 billion paid out...

should read

The $13.8 billion in cash transfers, paid out...

Page 13

The comparison of quarterly data reveals up to 0.4%...

should read

The comparison of quarterly data reveals deviations up to 4%...

Page 50

...on this data. This annual data shows...

should read

...on these data. These annual data show...

Page 70

Table 5.3

± (33,793)[1]

should read

± (33,793)[2]

Page 77

The extreme values, 371,655 from foreign sources and 207,420 from the taxation file as well as the 198,955...

should read

The extreme values, 371,655 from foreign sources and 207,420 from the taxation file as well as the 196,955...

Page 77

Table 5.8
Title

Comparison of Estimates of the Number of Emigrant Children Based on Income Tax Files

should read

Comparison of Estimates of the Number of Emigrant Children Based on Income Tax Files and Family Allowance Files

Page 77

Table 5.9
Footnote 2

Estimated as stated in table 5.6.

should read

Estimated as stated in table 5.7.

Page 81

t - h

should read

t - i

Page 83

Table 6.2
Footnote 2

...the paper by W. O'Hare (see footnote 2).

should read

...the paper by W. O'Hare (see footnote 2 page 81).

PREFACE

Population estimates are put to a wide variety of important uses, including those which arise from statutory requirements for the calculation of revenue transfers and cost-sharing programmes between the various levels of government. In these, as in other endeavours, timeliness and accuracy are of paramount importance.

The need for such information is met through the development of estimation techniques by Statistics Canada, taking into account the peculiarities of available data sources. These sources include administrative data files, developed to serve an entirely different purpose, but which contain valuable demographic information that can be exploited between censuses for the production of population estimates.

It is the Bureau's policy that the procedures used be open to public scrutiny, so that the strengths and weaknesses of the statistical products can be fully appraised. This volume brings together, under one cover, the methods used in the production of population and family estimates.

Ivan P. Fellegi,
Chief Statistician.

ACKNOWLEDGEMENTS

The population estimation methodology in this volume is an updated version of procedures developed over many years in the Bureau. Work pioneered by Nathan Keyfitz in the 1940s eventually led to the establishment of a population estimation programme, first under Allen LeNeveu, then under Margaret Fleming. Later, Yoshiko Kasahara laid the groundwork for estimating interprovincial migration from family allowance data. The programme received fresh impetus during the 1960s under the leadership of Karol Krotki, then Director of Research. Those who have more recently left their mark on the current state of methodology are John Kelly, former Chief of the Population Estimates Section, and Yolande Lavoie, Doug Norris, Normand Thibault and Claude Strohmenger, former staff members of Demography Division.

A major revision of the estimation methodology was initiated following the 1981 Census, in order to more completely integrate data from a number of sources (Census, Vital Statistics, and administrative data files from family allowance and taxation), and thus further improve timeliness and accuracy. This major undertaking was carried out by the current staff of the Population Estimates Section, including Rosemary Bender, Lise Champagne, Patricia Johnston, Gilbert Lagrange, Pierre Parent, Denis Prud'homme, Geoff Rowe, Edward Shin, and Ravi Verma, all under the leadership of Ronald Raby, Chief of the section, and K.G. Basavarajappa, Associate Director of Demography Division. The revised estimation methodology has also benefited from cross-fertilization with the provincial bureaus of statistics as well as with the Administrative Data Development, and Household Survey Methods Divisions, particularly its former Director, Richard Platek.

To put together a volume in a uniform style and format out of several background papers written by different authors at various periods, was no mean task. Every attempt was made to present the material in language understandable to a non-expert reader. Above all, there was a need to provide a focus for the volume around which its diverse subject matter could be organized. Although there are variations in the presentation, depending on the nature of the topic, three themes pervade the organization of each chapter — methods, data, and quality evaluation of the estimates. The job of coordinating the input and the editing of the volume was performed by A. Romaniuc. In his editorial function he was generously assisted by Terry Shevciw, Mali Jones, Robert Riordan and Lawrence Wise. While many colleagues in the Division were helpful in providing valuable assistance, Rosemary Bender, Ronald Raby and Ravi Verma deserve special recognition for their unfailing effort to ensure the successful completion of this work.

The development of the estimation methodology has benefited from the expert advice of a number of people outside the Bureau. Professor Roland Pressat of the Institut National d'Etudes Démographiques was consulted on the methodology for estimating families. Malcolm Britton from the Office of Population Censuses and Surveys, Great Britain, contributed to the methodology of internal migration estimates during his period as an exchange visitor at the Bureau. Professor Nathan Keyfitz of Harvard University offered insightful advice on a variety of estimation-related issues. Special recognition is due to Professor S.I. Mitra of Emory University, who spent part of his sabbatical year with Demography Division putting together an earlier version of this volume. Professor P. Krishnan from the University of Alberta and Richard A. Engels from the United States Census Bureau read and commented on an earlier version of the manuscript. While this valuable assistance by outside experts is gratefully acknowledged, the responsibility for any shortcomings in the publication rests with the authors.

The processing of the manuscript was handled by Ian Kisbee with the able assistance of Danielle St-Germain, Audrey Miles, and Dorothy Goyette, each of whom graciously performed the task of word processing.

Anatole Romaniuc,
Director,
Demography Division.

TABLE OF CONTENTS

TABLE OF CONTENTS - Continued

TABLE OF CONTENTS - Continued

LIST OF TABLES

TABLE OF CONTENTS - Continued

Page

LIST OF TABLES - Continued

OVERVIEW - Concluded

CHAPTER I

CHAPTER II

TABLE OF CONTENTS - Continued

Page

LIST OF TABLES - Continued

TABLE OF CONTENTS - Continued

Page

LIST OF TABLES - Continued

CHAPTER IV - Concluded

CHAPTER V

Page

LIST OF TABLES - Continued

TABLE OF CONTENTS - Concluded

Page

LIST OF TABLES - Concluded

LIST OF CHARTS

POPULATION AND FAMILY ESTIMATES: AN OVERVIEW

A long-standing policy at Statistics Canada, rearticulated in 1986[1], states that the methods used to produce the Bureau's statistical information should be open to public scrutiny. Doing so provides users with the opportunity to become acquainted with the strengths and limitations of the wide range of information available from the Bureau. This manual represents a direct response to that policy, by providing a comprehensive statement of the methods used to produce population estimates. These methods, while always available to the public, were heretofore scattered throughout various Statistics Canada publications and background papers. A reference to these sources, where applicable, is provided at the beginning of each chapter.

Scope of the Estimates Programme

The national census, conducted every five years, provides a wide range of demographic data on the Canadian population. Unlike some countries, however, Canada does not have a system of continuous population registration from which to derive basic demographic data on the state and movement of the population for *non-census* years. To fill this gap in the information system, Statistics Canada has developed over the years since the 1940s, a programme of population and family estimates.

Specifically, the estimates programme encompasses the following six areas:

(1) Total population of Canada, provinces and territories, census divisions and census metropolitan areas;

(2) Population by age, sex and marital status, for Canada, provinces and territories;

(3) Number of families, their size, children's age and parents' age and sex;

(4) Internal migration;

(5) International migration;

(6) Special requests for demographic estimates at various geographic levels (e.g., for economic regions, municipalities, etc.).

---

[1] *Policy on Informing Users of Data Quality and Methodology*, Statistics Canada, Ottawa, March 1986.

Use of Estimates

The population estimates produced by Statistics Canada have wide ranging applications in the areas of planning, marketing and programme evaluation in both the public and private sectors. Table 1 gives an indication of the volume of demand for population estimates. Some specific uses to which the estimates are put include:

(1) Calculation of social and economic indicators (vital rates, unemployment rates, school enrollment rates, etc.) in which the population, or a part thereof, serves as the denominator;

(2) Calculation of weights for use in Statistics Canada's Surveys (Labour Force Survey, Household Facilities Survey, etc.);

(3) Preparation of population projections by Statistics Canada, where estimates of population by age and sex are used as the base population;

(4) Determination by the Government of Canada of the annual level of immigration; and

(5) Calculation of revenue transfers and grants under various statutory programmes, as well as cost-sharing agreements between federal, provincial and municipal governments.

TABLE 1. Requests and Publication Sales for Population Estimates, 1985

Estimate	Frequency of Request			Publications	
	Total	CANSIM [1] (No. of accesses)	User Services [2] and Demography Division	Catalogue No.	Copies Sold
Population Estimates for Canada, Provinces and Territories	**44,421**	38,916	5,505	91-210 91-001	1,200 3,600
Sub-provincial Estimates	**1,451**	-	1,451	91-211 & 91-212	1,600
Other Estimates	**1,922**	-	1,922	N.A.	N.A.
Total	**47,794**	**38,916**	**8,878**		**6,400**

N.A.: Not applicable
-: Available only as of 1986
1 CANSIM: Canadian Socio-Economic Information Management System.
2 Requests received in all Statistics Canada Regional Offices and in the Central Inquiries Service in Ottawa.

This last application adds an important dimension to the population estimates programme. Statistics Canada is under statutory obligation to supply annual population figures, certified by the Chief Statistician, in order to establish the final amounts payable under the Federal-Provincial Fiscal Arrangements and Federal Post-Secondary Education and Health Contributions Act of 1977. The $13.8 billion paid out under this act is by far the largest amount in which population estimates play a role. Since population is one of the key variables in the allocation formulae under this act, timely estimates with a high degree of accuracy are essential.[2] Table 2 identifies a few selected programmes, along with the amount transferred during the 1985-86 fiscal year.

Types of Estimates

Estimates can be categorized in three ways, the first of which refers to the frequency with which they are issued. In this regard, there are quarterly and annual postcensal estimates, the former being issued for periods ending on the first day of January, April, July and October, while annual estimates are referenced to June 1st, to coincide with the date on which the Canadian census is taken (i.e., June 1 or very close to it). All other estimate configurations — age, sex, marital status and families, as well as estimates of total population of census divisions and census metropolitan areas — are produced on an annual basis.

The second classification involves the degree of finalization of the estimates. Specifically, at the national and provincial levels, three types of postcensal estimates are issued: (1) preliminary; (2) updated, and; (3) final. For subprovincial areas, only preliminary and final estimates are produced. Reference dates and the time-lag associated with the production of the three types of estimates are shown in Table 3, indicating that preliminary estimates are available within three to four months of the reference date, while final estimates take twenty months to complete. The production of three successive series of estimates, while sometimes criticized on the grounds that it confuses the user, is the strategy that best satisfies the need for both timeliness and accuracy.

Finally, intercensal estimates are generated as soon as census population counts become available. This third type involves retrospective adjustment to bring the postcensal estimates in line with the census counts.

Estimation Methods and Data Sources

Postcensal estimates of national population at a given point in time are obtained by adding the number of births and immigrants to, and by subtracting

---

[2] I.P. Fellegi, "Should the Census Count be Adjusted for Fiscal Allocation Purposes? — Some Statistical Considerations"; A. Romaniuc and R. Raby, "The Impact of Census Underenumeration on Selected Federal-Provincial Transfer Payments", in *Underenumeration in Canadian Censuses and Its Ramifications: An Examination of Selected Aspects,* Statistics Canada, Mimeograph, Ottawa, June 1980.

TABLE 2. Selected Federal Cash Transfers to the Provinces, Territories and Municipalities, Requiring Population Estimates, Fiscal Year 1985-86

Selected Programmes and Departments	Dollars (in millions)	Type of Population Data Used in Calculations
Programmes Administered Under the Federal-Provincial Fiscal Arrangements and Federal Post-Secondary Education and Health Contributions Act, 1977		
Fiscal Equalization Programme (Finance)	5,019.0	Total Population (Annual and Quarterly)
Insured Health Services Programme (Health and Welfare)	5,331.9	Total Population (Annual and Quarterly)
Extended Health Services (Health and Welfare)	1,075.4	Total Population (Annual and Quarterly)
Post-secondary Education Financing (Finance - Secretary of State)	2,365.5	Total Population (Annual and Quarterly)
Subtotal	**13,791.8**	
Other Programmes		
Canada Student Loans Plan (Secretary of State)	592.9	Population (18-24) August 1
Local Employment Assistance and Development - LEAD (Employment and Immigration)	90.8	Total Population for Communities Under 50,000
Assistance to Provinces for the Provision of Legal Aid in Matters of Criminal Law	38.7	Total Population (Annual)
Statutory Sudsidies (Finance)	35.8	Total Population, June 1
Emergency Planning (Emergency Planning Canada)	9.1	Total Population (Annual)
Assistance to Provinces for the Provision of Compensation to Victims of Violent Crime	2.6	Total Population (Annual)
University Research Programme (Communications)	1.1	Based on Regional Demographic Factors
TOTAL	**14,562.8**	

Source: Federal-Provincial Programmes and Activities, 1985-86, Reference No. ISSN 0823-9193.

TABLE 3. Types of Postcensal Population Estimates

Estimate	Geographic Areas	Reference Date	Finality Status and Time-Lag after Reference Date	Publication (Catalogue No.)
1. Total Population	Canada, Provinces and Territories	January 1 April 1 July 1 October 1	Preliminary (3-month lag) Updated (12-month lag) Final (20-month lag)	No. 91-001, (Quarterly)
		June 1	Preliminary (3-month lag) Updated (12-month lag) Final (20-month lag)	No. 91-210, (Annual)
	Census Divisions and Census Metropolitan Areas	June 1	Preliminary (4-month lag) Final (20-month lag)	No. 91-211, (Annual) No. 91-212, (Annual)
2. Age, Sex and Marital Status	Canada, Provinces and Territories	June 1	Preliminary (4-month lag) Updated (12-month lag) Final (20-month lag)	No. 91-210, (Annual) Available on request Available on request
3. Number of Census Families	Canada, Provinces and Territories	June 1	Preliminary (4-month lag) Updated (12-month lag) Final (20-month lag)	No. 91-204, (Annual)

the number of deaths and emigrants from, the most recent census population. For subnational estimates, numbers of internal in-migrants, as well as out-migrants, are also required. Known as the "Component Method", the process is straightforward in itself, but has required many methodological enhancements in order to improve on the dual requisites of timeliness and accuracy.

In the absence of readily available data, special procedures had to be devised for estimating marital status and family information. In this case, relevant parameters were derived by a method of synthetic estimation from proxy statistical information. The development of methods for estimating interprovincial migration and international emigration from family allowance and personal income tax files stands out as a significant methodological innovation, as does the method devised to generate preliminary estimates of population for census divisions and census metropolitan areas. The latter was accomplished by the development of a regression model using symptomatic indicators derived from locally available data including medicare, hydro-connections, and family allowance. Table 4 identifies the methods and data sources used for various types of postcensal estimates.

Outline of the Manual

In general, the overall organization of the manual follows a pattern dictated by the movement from aggregate estimates into the various disaggregations for which estimates are provided. Since disaggregation is possible only through the application of increasingly numerous and more complex techniques, the manual can also be seen as following a path towards increasing complexity of method.

Individual chapters are organized around three recurring themes: (1) method; (2) data sources, and; (3) quality evaluation. In this regard, an attempt is made in each chapter to answer the corresponding questions: (1) how are the estimates produced; (2) what are the data and their sources, and; (3) how reliable are the estimates in light of various validation criteria? In this way, the reader is provided with a clear and succinct picture of the strengths and weaknesses of each procedure.

The manual begins with the presentation of the methods used to produce postcensal estimates of the total population at the national, provincial levels. This is followed by the presentation, in Chapter II, of the methods used to generate the postcensal estimates of the population disaggregated by age, sex and marital status, again at the national and provincial levels. Chapter III deals with the intercensal estimates for the same populations.

Estimates of interprovincial migration play a crucial role in the process of population estimation at several levels of aggregation. This importance, along with the unconventionality of the data types used in their production (family allowance and income tax data), warrants the lengthy and elaborate treatment

TABLE 4. Data Sources and Methods Used

Method Used	Estimates	Data Sources
(A) Component Method Uses the census counts plus the effect of events such as births, deaths, migrations, marriages, divorces and new widowhoods which generate the change in total population or age, sex and marital status distribution of the population	1. Total Population: for Canada, Provinces and Territories, Census Divisions and Census Metropolitan Areas	- Census Counts (Statistics Canada) - Births, Deaths (Statistics Canada) - Immigration (EIC) - Child Interprovincial Migration and Emigration from Family Allowance (HWC) - Interprovincial and Intra-provincial Migration from Income Tax Files (Revenue Canada)
	2. Age, Sex and Marital Status for Canada, Provinces and Territories	- Census Counts/Population Estimates (Statistics Canada) - Births, Deaths, Marriages, Divorces, New Widowhood (Statistics Canada) - Child Interprovincial Migration and Emigration from Family Allowance (HWC) - Immigration (EIC) - Emigration from Canada to the U.S.A. (U.S. Dept. of Immigration) - Interprovincial Migration from Income Tax Files (Revenue Canada)
	3. Age and Sex for Census Divisions and Census Metropolitan Areas	- Census Counts and Mobility (Statistics Canada) - Births, Deaths (Statistics Canada) - Child Interprovincial Migration and Emigration from Family Allowance Files (HWC) - Health Care Recipients (Bureau of Statistics, Saskatchewan and Alberta) - Interprovincial Migration from Income Tax Files (Revenue Canada)
	4. Number of Census Families for Canada, Provinces and Territories	- Census Counts of Families - Deaths, Marriages, Divorces (Statistics Canada) - Immigration (EIC) - Regular Family Allowance Accounts (HWC) - Population Estimates (Statistics Canada)

TABLE 4. Data Sources and Methods Used - Concluded

Method Used	Estimates	Data Sources
(B) Regression Method Uses the ratio or difference correlation which relates the changes in the proportional shares of the population with the symptomatic indicators	Total Population for Census Divisions and Census Metropolitan Areas	- Census Counts (Statistics Canada) - Number of Children Receiving Family Allowance (HWC) - Health Care Recipients (Bureau of Statistics, Saskatchewan and Alberta)

Note: EIC: Employment and Immigration Canada
HWC: Health and Welfare Canada

given in Chapter IV. Chapter V deals with the methods used to produce estimates of emigration, using the same data types as those for interprovincial migration.

Estimates of local area population — census divisions and census metropolitan areas — are dealt with in Chapter VI, where an explanation of the regression method, designed to produce preliminary estimates from symptomatic indicators, accounts for the major part of the chapter. The manual ends with the exposition, in Chapter VII, of the rather elaborate method used to produce the family estimates.

For the benefit of those who are not fully conversant with the subject matter of this manual, a glossary of principal terms is included in Appendix I. A list of mathematical symbols is presented in Appendix II. An attempt was made to simplify and standardize the notation system in the formulae throughout this manual. As such, notations in the manual may differ from those used in previous publications of population estimates.

CHAPTER I

POSTCENSAL ESTIMATES: TOTAL POPULATION*

The material in this chapter, as in most of the volume, is organized around the three topics: methodology, data sources and evaluation. Accordingly, the chapter begins with the presentation of the estimation algorithm — the tool with which postcensal estimates of the total population are generated. Subsequent discussion centres on the data used, and includes an assessment of their strengths and limitations. The chapter ends with an evaluation of the quality of the estimates in terms of "error of closure" — a procedure which provides a measure of the accuracy of the postcensal estimates against census counts.

Estimation Methodology

Estimates of the population of Canada, and of its provinces and territories, are produced by the component method. Starting with the census population (time t) as the base, the estimate of population for any reference date (t + i), according to this method, is the change in population size during the period (t, t + i), added to the base population. Estimates of population are first produced for each province and territory, and then summed to obtain an estimate of the population of Canada for the same reference date.

Population growth over a given period is the net result of changes in the components. Births and immigrants from other countries are the incremental components, while deaths and emigrants to other countries are the decremental components. Net interprovincial migration can be an increment or a decrement, depending on the sign of the value.

Symbolically, the method is expressed as follows:

$$\hat{P}_{t+i} = P_t + [B_{(t,\,t+i)} - D_{(t,\,t+i)} + I_{(t,\,t+i)} - E_{(t,\,t+i)}] + N_{(t,\,t+i)} \qquad (1)$$

where, for any given province:

$\hat{P}_{(t+i)}$	= estimate of population at time t + i;
P_t	= census population counts at time t;
B	= number of births between time t and t + i;
D	= number of deaths between time t and t + i;
I	= number of immigrants between time t and t + i;

* This chapter is based mainly on a background paper prepared by Ronald Raby.

E = number of emigrants between time t and t + i;

N = number of net interprovincial migrants between time t and t + i; and

(t, t + i) = interval between the last census date and the reference date of the estimate.

For June estimates, t, t + i refers to the period from June 1 of the census year (t) to May 31 of the i^{th} year (t + i). For estimates produced as of the first day of January, April, July and October, t, t + i covers the period from June 1 of the census year (t) to December 31, March 31, June 30 and September 30 of the i^{th} year (t + i), respectively.

Preliminary estimates, produced as soon as all data become available, are generally released within three months of the reference date. As these data are refined, the estimates are revised accordingly. Revision of estimates is also undertaken when estimates with prior reference dates are finalized. The final estimates are completed some 20 months after the end of the period to which they apply.

Data Sources

Monthly data are used for most of the components of population change. The following identifies the sources and discusses the quality of the data used for the calculation of the postcensal estimates of population (see Table 1.1).

Base Population

The base population is taken from the Census of Canada. The universe of the *de jure*[1] censuses, which have been conducted in Canada every five years since 1951, extends to the following persons:

- all persons whose usual place of residence is somewhere in Canada;
- Canadian government employees and their families stationed abroad;
- members of the Canadian Armed Forces and their families stationed abroad; and
- crews of Canadian merchant vessels.

Not included in census counts are:

- government representatives of other countries (and their families) attached to the legation, embassy or other diplomatic body, residing in Canada;
- members (and their families) of the armed forces of other countries, stationed in Canada;

[1] A *de jure* census is designed to count persons according to their usual place of residence.

- students attending school in Canada whose usual residence is outside Canada;
- residents of another country visiting Canada temporarily; and
- workers from another country staying in Canada temporarily.

TABLE 1.1. Data Sources Used in Producing Postcensal Population Estimates for Canada, Provinces and Territories

Data	Sources	References
Base Population	Statistics Canada (Census of Canada)	Catalogue No. 92-901, 1981 Census
Births	Statistics Canada (Vital Statistics Section, Health Division)	Catalogue Nos. 84-204, Annual, and 84-001, Quarterly
Deaths	Statistics Canada (Vital Statistics Section, Health Division)	Catalogue Nos. 84-204, Annual, and 84-001, Quarterly
Immigration	Employment and Immigration Canada (Planning and Programme Management Branch)	Monthly Statistics
Emigration	1. Children (0-17 years): Health and Welfare Canada (Planning and Evaluation, Income Securities Programmes)	Family Allowance File M0024 based on change of address notifications, obtained every six months
	2. Adults: Estimates derived by Statistics Canada (Demography Division)	The estimation methodology (see Chapter V) also makes use of data obtained from Revenue Canada Income Tax Files
Interprovincial Migration	**Final Data**	
	1. All ages: Revenue Canada Income Tax files	Annual data are distributed quarterly using Family Allowances File M0024
	Preliminary Data	
	1. Children (0-17 years): Health and Welfare Canada (Planning and Evaluation, Income Securities Programmes)	Special Monthly Family Allowance tables based on change of address notifications
	2. Adults: Estimates derived by Statistics Canada (Demography Division)	The estimation methodology (see Chapter IV) also makes use of data obtained from Revenue Canada Income Tax Files

Recent censuses have proven to be of exceptional quality[2], particularly in terms of completeness of enumeration. A special procedure, the Reverse Record Check (RRC), has been developed to determine the number and characteristics of individuals and households missed in the census. The RRC is a procedure involving tracing and checking, based on a probabilistic sample of individuals drawn from four different sample frames. These frames are: (1) records from the previous census; (2) Vital Statistics records on births during the last intercensal period; (3) information from Employment and Immigration Canada on immigrants during the intercensal period, and; (4) persons identified as missed in the previous census. In 1981, the sample size from these four frames totaled 36,423 persons. All of these individuals were subject to tracing and follow-up procedures, culminating in the classification of each as "enumerated", "missed", "deceased", "emigrated" or "tracing failed". The Reverse Record Check for the last three censuses has revealed an undercoverage rate in the order of 2 % for the whole of Canada, with variations among provinces ranging between 1 % and 3 % in 1981.

TABLE 1.2. Estimated Population Undercoverage Rates in the 1971, 1976 and 1981 Censuses: Canada and Provinces

Geographic Area	1971 Census		1976 Census		1981 Census	
	Estimated Rate	Standard Error	Estimated Rate	Standard Error	Estimated Rate	Standard Error
			Percent			
Canada[1]	**1.93**	**0.09**	**2.04**	**0.10**	**2.01**	**0.09**
Newfoundland	2.25*	0.72	1.10*	0.39	1.74*	0.45
Prince Edward Island	1.23*	1.13	0.38*	0.25	1.17*	0.54
Nova Scotia	1.33*	0.45	2.16	0.37	1.05*	0.34
New Brunswick	1.65*	0.56	0.86	0.34	1.81	0.30
Quebec	2.10	0.19	2.95	0.25	1.91	0.21
Ontario	1.68	0.12	1.52	0.17	1.94	0.14
Manitoba	1.13*	0.38	1.07*	0.33	0.98*	0.35
Saskatchewan	1.00*	0.37	1.33*	0.34	0.99*	0.37
Alberta	2.55	0.44	1.49	0.26	2.54	0.36
British Columbia	2.89	0.39	3.13	0.31	3.16	0.33

* Indicates that the estimate has a very large standard error in relation to its size (over 20%).
[1] Excluding the Yukon and the Northwest Territories.
Source: 1971, 1976 - Statistics Canada, Catalogue No. 99-840, Table 16, and unpublished data. 1981 - Statistics Canada, Catalogue No. 99-904, (forthcoming).

[2] See: *Quality of Data, Series 1: Sources of Error-Coverage*, Catalogue 99-840, 1976 Census of Canada, Statistics Canada, Ottawa; *Response Error in the 1976 Census of Population and Housing*, Catalogue 8-2400-531, Working Paper by K.P. Krotki, Statistics Canada, Census Survey Methods, Ottawa 1980; *Data Quality - Total Population*, Catalogue 99-904, Statistics Canada, Ottawa (forthcoming).

Births and Deaths

Births and deaths have been recorded on a regular basis since 1921. Because the recording of vital events is required by law, the data should not be deficient in either coverage or quality. There are, however, some disparities in the data, due to delays in the recording of events past the annual cut-off date of April 30th.

Because preliminary data on births and deaths are affected by delays in the receipt of registrations by provincial offices, as well as by carry-overs of data from previous quarters, preliminary estimates of vital events are obtained by means of a multiple regression model based on the number of registrations filed provincially in each quarter. An indication of the quality of these estimates, both quarterly and annually, is provided in Tables 1.3 and 1.4, in which the estimated and final data on births and deaths are compared for 1984. These comparisons indicate that the preliminary data for 1984 overstated the actual number of births and deaths in Canada by 1 % and 2 % respectively, but at the provincial level, excluding the territories, the percentage deviations were in some cases as high as 4 %. The comparison of quarterly data reveals deviations up to 4 % for births in the January-March quarter, as well as for deaths in the October-December quarter for Canada. Deviations based on quarterly data were as high as 21 % at the provincial level (excluding the territories).

It should be noted that the Vital Statistics universe differs slightly from that of the Census, in that the former includes births and deaths among alien students and workers (and their families) temporarily in Canada, and births and deaths among Canadian citizens and landed immigrants working or studying in other countries or visiting abroad for more than a year. These categories are not targeted for coverage by the census.

Immigration

All immigrants are required to submit an "Immigrant Visa and Record of Landing" form upon their arrival in Canada. Immigration statistics, therefore, include only those immigrants who are lawfully admitted into Canada to establish permanent residence. They do not include immigrants entering Canada illegally; returning Canadian citizens who worked, studied or visited abroad; or visitors, boarders, commuters, students, workers, diplomatic and consular representatives and their families from other countries.

Except for a very small percentage, (0.27 % in 1982 and 0.02 % in 1983) the file identifies the province in which the immigrant intends to settle. There may be some discrepancies between the intended destinations stated on the form and the actual province of settlement. Due to insufficient information, however, the magnitude of these discrepancies cannot be assessed.

Immigration data are produced in a preliminary form, and in a final form, the latter providing detailed data by month of arrival. These final statistics

TABLE 1.3. Ratio of Final to Preliminary Data on Births for 1984: Canada, Provinces and Territories

Geographic Area	January-March	April-June	July-September	October-December	January-December
Canada	**0.96**	**0.99**	**1.01**	**0.99**	**0.99**
Newfoundland	1.08	1.03	1.02	0.94	1.02
Prince Edward Island	1.00	0.97	1.02	1.02	1.00
Nova Scotia	0.99	0.98	1.00	0.98	0.99
New Brunswick	0.89	0.97	1.00	0.99	0.96
Quebec	0.90	0.99	0.99	0.97	0.96
Ontario	0.95	1.02	1.06	1.00	1.01
Manitoba	0.91	0.97	1.01	1.01	0.97
Saskatchewan	1.21	0.96	1.04	0.97	1.04
Alberta	1.06	0.96	0.96	0.98	0.99
British Columbia	0.98	0.97	0.97	0.98	0.98
Yukon	1.09	1.02	0.88	1.13	1.02
Northwest Territories	1.50	1.38	0.93	1.19	1.21

TABLE 1.4. Ratio of Final to Preliminary Data on Deaths for 1984: Canada, Provinces and Territories

Geographic Area	January-March	April-June	July-September	October-December	January-December
Canada	**1.02**	**1.01**	**0.98**	**0.96**	**0.98**
Newfoundland	0.99	1.17	1.07	0.83	1.01
Prince Edward Island	1.13	1.03	0.88	0.99	1.03
Nova Scotia	1.18	1.05	0.92	0.95	1.00
New Brunswick	0.94	1.02	0.94	1.00	0.98
Quebec	0.94	0.98	0.96	1.00	0.96
Ontario	1.02	1.01	1.03	0.94	0.99
Manitoba	1.00	1.02	0.92	1.00	0.99
Saskatchewan	0.98	0.97	0.97	0.94	0.97
Alberta	1.19	1.02	0.93	0.95	0.98
British Columbia	1.10	1.04	1.00	0.93	1.01
Yukon	0.83	1.15	0.97	0.78	0.90
Northwest Territories	0.96	1.63	1.07	1.20	1.25

Source: Statistics Canada, Catalogue Nos. 84-001 and 84-204.

TABLE 1.5. Ratio of Final to Preliminary Data on Immigrants for 1984: Canada, Provinces and Territories

Geographic Area	January-March	April-June	July-September	October-December	January-December
Canada	**1.04**	**0.98**	**0.99**	**1.00**	**1.00**
Newfoundland	1.00	0.98	1.02	1.07	1.01
Prince Edward Island	1.06	0.95	1.03	1.04	1.01
Nova Scotia	1.05	0.96	1.02	1.00	1.00
New Brunswick	1.03	0.98	1.00	0.98	1.00
Quebec	1.04	1.01	0.97	1.00	1.00
Ontario	1.04	0.98	0.99	1.00	1.00
Manitoba	1.02	0.96	1.02	1.00	1.00
Saskatchewan	1.01	0.97	1.02	1.00	1.00
Alberta	1.04	0.99	0.98	1.00	1.00
British Columbia	1.03	0.97	1.01	1.00	1.00
Yukon	1.00	1.00	1.00	1.00	1.00
Northwest Territories	1.07	1.15	0.83	1.00	1.00

Source: Employment and Immigration Canada, Special Tabulations.

do not become available until approximately 8 months after the end of the calendar year. The immigration data used in the preliminary population estimates are produced by month of compilation rather than by the month in which the immigrants actually arrived. Nevertheless, preliminary estimates of immigration by calendar year are very close to the final figures (Table 1.5). With respect to the quarterly estimates, however, the differences between the preliminary and the final figures are greater. The deviations in 1984 were as high as 4 % for Canada and 7 % for the provinces, excluding the territories.

Emigration and Interprovincial Migration

In Canada there is no provision for recording interprovincial migration or emigration, and, as such, both of these components require estimation. This is done on the basis of data derived from personal income tax files, and from family allowance files. As these two components require more extensive treatment, the description of the methodology and data used in their estimates are dealt with in chapters IV and V, respectively.

Evaluation of Estimates

Strictly speaking, there are no objective criteria for validating the postcensal estimates[3]. The Census, however, is considered to be a reliable benchmark. The difference between the two, called the "error of closure", provides a measure of accuracy for the postcensal estimates.

In this section, a comparison is first made between the consecutive sets of estimates — preliminary, updated and final, for the years 1983 and 1984. Second, the final estimates for the intercensal period are gauged against the census counts, using error of closure.

Comparison of the Preliminary, Updated and Final Estimates

The percentage differences between the three successive sets of annual postcensal estimates (preliminary, updated and final) are presented in Table 1.6. These estimates show very small differences (less than one-half of one percent) for the provinces. Quarterly estimates display somewhat larger variations, but are still under 1 % for most provinces. Given their closeness, why should all of these successive estimates be produced? The answer to this question is that, while the preliminary estimates would probably be quite acceptable for many purposes, even very small understatements or overstatements of the population may entail considerable fund misallocations under cost-sharing programmes. Furthermore, while minimal at the aggregate level, differences between preliminary and final estimates may be quite significant at various sub-aggregate levels such as particular age groups or marital statuses. Hence, in spite of the possible inconvenience to some users, who face a choice involving three estimates with different time-lags, their production satisfies the need dictated by the requirements for timeliness, accuracy, and consistency.

Comparison of Estimates with Census Counts

The error of closure is defined as the difference between the enumerated census population and the estimated population, for any disaggregated group, or for any summation of such disaggregations up to and including the total population. The error of closure is calculated as follows:

$$\varepsilon = \hat{P} - P \qquad (2)$$

where: ε = error of closure;

$\hat{P}$ = estimated population;

P = census population.

[3] None of the Atlantic provinces, nor the Yukon, produce independent population estimates. The remaining provinces and the Northwest Territories do produce regional estimates, but use different methodologies and/or data bases (such as provincial health insurance files, hydro connections, driver's licence files, etc.). Moreover, some provinces adjust their final estimates in accord with those produced by Statistics Canada.

TABLE 1.6. Percent Differences Between Preliminary, Updated and Final Postcensal Estimates: Canada, Provinces and Territories, June 1, 1983 and 1984

Geographic Area	Percent Differences					
	Preliminary and Updated		Updated and Final		Preliminary and Final	
	1983	1984	1983	1984	1983	1984
Canada	**0.02**	**0.02**	**0.01**	**0.00**	**0.02**	**0.02**
Newfoundland	0.30	0.09	0.03	−0.05	0.33	0.05
Prince Edward Island	0.08	−0.27	−0.32	−0.19	−0.24	−0.46
Nova Scotia	0.00	−0.25	−0.26	0.12	−0.26	−0.14
New Brunswick	0.01	−0.12	−0.16	−0.05	−0.14	−0.18
Quebec	0.10	0.12	0.07	−0.05	0.18	0.06
Ontario	0.00	−0.11	−0.10	0.05	−0.11	−0.06
Manitoba	0.09	−0.15	−0.17	−0.07	−0.09	−0.22
Saskatchewan	0.06	−0.26	−0.15	0.19	−0.09	−0.06
Alberta	−0.16	0.35	0.31	−0.41	0.15	−0.06
British Columbia	−0.07	0.19	0.19	0.25	0.12	−0.45
Yukon	0.45	−1.21	−0.89	−1.52	−0.45	−2.71
Northwest Territories	−0.41	−0.30	0.21	−0.30	−0.21	−0.61
Mean Absolute Error	**0.14**	**0.29**	**0.24**	**0.27**	**0.20**	**0.42**
Weighted[1] Mean Absolute Error	**0.06**	**0.16**	**0.13**	**0.12**	**0.14**	**0.12**

[1] Weights based on provincial populations.

The total error represented by ε has two sources: census under or over-coverage, and component error. Both types of error have unique sources, depending on the type of estimate under consideration. In any case, it is difficult, if not impossible, to determine the contribution of each to the total error of closure.

Table 1.7 presents the differences (errors of closure) between the census counts and postcensal estimates. At the national level, the differences are small (0.5 % for 1971, 0.6 % for 1976 and −0.3 % for 1981). At the provincial level, however, the differences are understandably larger, since the provincial estimates are affected by errors in estimating interprovincial migration. Nevertheless, excluding the territories, the provincial postcensal estimates fall within

TABLE 1.7. Error of Closure: Canada, Provinces and Territories, June 1, 1971, 1976 and 1981

Geographic Area	Population	Error of Closure	Percent Error[1]
	1971		
Canada	**21,568,311**	**109,254**	**0.51**
Newfoundland	522,104	1,688	0.32
Prince Edward Island	111,641	−844	−0.76
Nova Scotia	788,960	−19,317	−2.45
New Brunswick	634,557	−2,762	−0.44
Quebec	6,027,764	4,554	0.08
Ontario	7,703,106	108,918	1.41
Manitoba	988,247	−133	−0.01
Saskatchewan	926,242	1,921	0.21
Alberta	1,627,874	5,120	0.31
British Columbia	2,184,621	10,235	0.47
Yukon	18,388	−1,220	−6.63
Northwest Territories	34,807	1,094	3.14
Mean Absolute Error[3]			**1.35**
Weighted[4] Mean Absolute Error			**0.71**
	1976		
Canada	**22,992,604**	**133,209**	**0.58**
Newfoundland	557,725	−1,059	−0.19
Prince Edward Island	118,229	1,863	1.58
Nova Scotia	828,571	7,699	0.93
New Brunswick	677,250	10,237	1.51
Quebec	6,234,445	6,297	0.10
Ontario	8,264,465	88,321	1.07
Manitoba	1,021,506	12,396	1.21
Saskatchewan	921,323	8,363	0.91
Alberta	1,838,037	−1,745	−0.09
British Columbia	2,466,608	1,737	0.07
Yukon	21,836	−510	−2.34
Northwest Territories	42,609	−390	−0.92
Mean Absolute Error[3]			**0.91**
Weighted[4] Mean Absolute Error			**0.61**

See footnote(s) at the end of table.

TABLE 1.7. Error of Closure: Canada, Provinces and Territories, June 1, 1971, 1976 and 1981 - Concluded

Geographic Area	1981		
	Population [2]	Error of Closure	Percent Error[1]
Canada	**24,341,701**	**−61,083**	**−0.25**
Newfoundland	567,670	7,101	1.25
Prince Edward Island	122,501	−380	−0.31
Nova Scotia	847,421	−240	−0.03
New Brunswick	696,382	−1,951	−0.28
Quebec	6,438,180	−37,643	−0.58
Ontario	8,624,712	31,492	0.37
Manitoba	1,026,236	8,561	0.83
Saskatchewan	968,262	−5,041	−0.52
Alberta	2,237,286	−53,886	−2.41
British Columbia	2,744,163	−6,048	−0.22
Yukon	23,151	−488	−2.11
Northwest Territories	45,737	−2,560	−5.60
Mean Absolute Error [3]			**1.21**
Weighted [4] Mean Absolute Error			**0.64**

[1] $\frac{\text{Estimate - Census}}{\text{Census}} \times 100$

[2] The June 3, 1981 Census figures were adjusted to refer to June 1, 1981.

[3] Mean absolute error is the sum of the absolute values of the percent differences divided by the number of categories.

[4] Weights based on provincial populations.

1 % of the census counts with few exceptions (Nova Scotia at −2.5 % in 1971, and Alberta at −2.4 % in 1981). It should be noted that errors of closure represent errors that have accumulated over the five-year period since the previous census.

As discussed in the section on data sources, each of the event components carries a degree of bias and error. Generally speaking, however, the data on births, deaths, and even immigration can be regarded as being fairly accurate, while the estimates of interprovincial migration are less so. The most prone to error are the estimates of emigration, but the impact of emigration on population change is smaller than that of any of the other components. The contribution of censuses to the error of closure is due predominantly to census undercoverage,[4] which can be measured by means of the Reverse Record Check, explained earlier.

Adjustment for undercoverage, however, does not necessarily reduce the error of closure, as shown in Table 1.8. In fact, a test for Canada and provinces indicated that the error of closure increased slightly after adjustment, from −0.25 % to −0.33 %. In some provinces the differences between estimates and the 1981 Census counts are attenuated when adjusted for undercoverage, while in others they are widened. Nevertheless, from the foregoing analysis, it can be concluded that the postcensal estimates for Canada and provinces are of high quality. Their accuracy, as measured against the census counts, is remarkable.

TABLE 1.8. Error of Closure, With and Without Census Undercoverage Adjustment: Canada, Provinces and Territories, June 1, 1981

Geographic Area	With Adjustment		Without Adjustment	
	Number	Percent	Number	Percent
Canada	**−81,717**	**−0.33**	**−61,083**	**−0.25**
Newfoundland	3,262	0.56	7,101	1.25
Prince Edward Island	−1,391	−1.12	−380	−0.31
Nova Scotia	−2,059	−0.24	−240	−0.03
New Brunswick	145	0.02	−1,951	−0.28
Quebec	26,832	0.41	−37,643	−0.58
Ontario	−12,363	−0.14	31,492	0.37
Manitoba	9,438	0.91	8,561	0.83
Saskatchewan	−2,313	−0.24	−5,041	−0.52
Alberta	−84,431	−3.68	−53,886	−2.41
British Columbia	−15,718	−0.55	−6,048	−0.22
Yukon [1]	−508	−2.15	−488	−2.11
Northwest Territories [1]	−2,611	−5.59	−2,560	−5.60
Mean Absolute [2] Error		**1.30**		**1.21**
Weighted Mean [3] Absolute Error		**0.65**		**0.64**

[1] Estimated assuming the same undercoverage rates as Canada excluding the territories.

[2] Mean absolute error is the sum of the absolute values of the percent differences divided by the number of categories.

[3] Weighted mean absolute error is the sum of the absolute error for each province weighted by the relative size of the provincial population.

[4] In July 1981, after extensive study and consultation with the federal departments concerned, the provincial statistical bureaus, and experts, the Chief Statistician concluded that the 1981 Census of Canada should not be corrected for undercoverage.
The relevant background documents are obtainable from Demography Division, Statistics Canada.

CHAPTER II

POSTCENSAL POPULATION ESTIMATES BY SEX, AGE AND MARITAL STATUS*

Postcensal estimates of population by age, sex and marital status are produced using the cohort component method. This method is similar to that described in Chapter I used to derive total population estimates. Certain modifications must be made to it, however, in order to produce these more detailed estimates. This chapter presents and discusses the modified component method used to produce population estimates disaggregated by age, sex and marital status.

Estimation Methodology

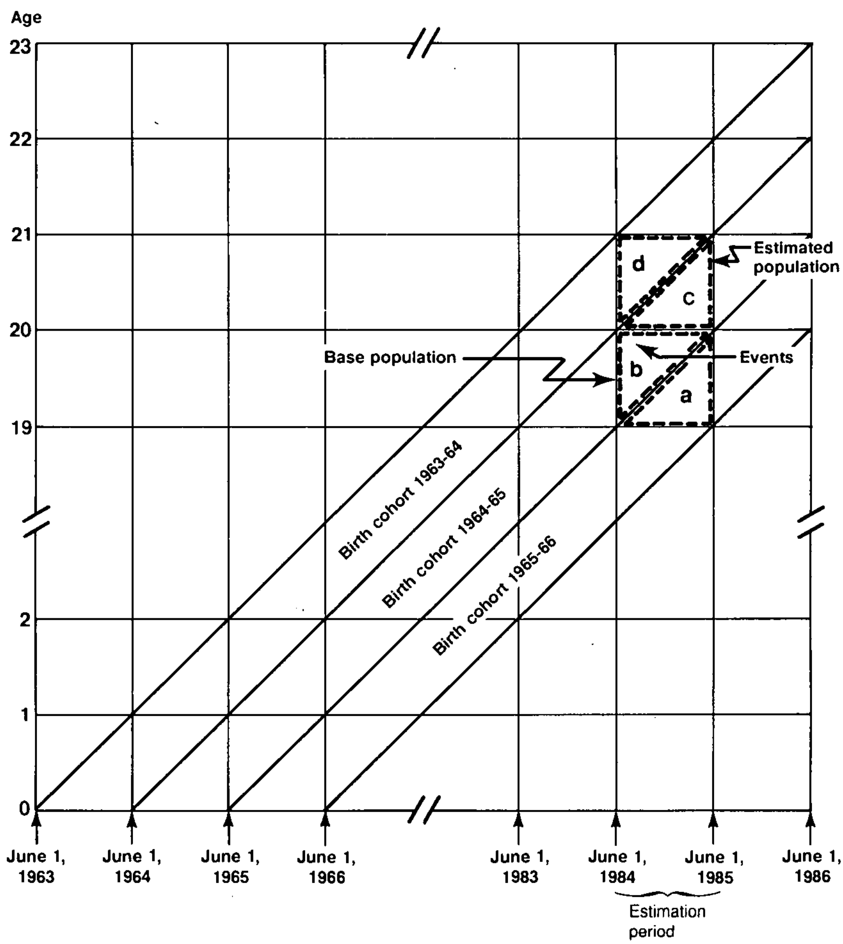
Use of the component method for estimating population by age, sex and/or marital status poses particular problems because of the nature of the disaggregation. This can be illustrated by means of the Lexis diagram, Chart 2.1, displaying the dual (cohort and period) classifications of the demographic events experienced by the population. Take, for example, those aged 19 as of June 1, 1984, who belong to the cohort born between June 1, 1964 and May 31, 1965. The demographic events experienced by this group during the estimation period are represented by triangles "b" and "c". In the Canadian registration system (Vital Statistics), however, the event data are recorded by age at the time of the event. Thus, events occurring to those aged 19 years during the estimation period (June 1, 1984 to June 1, 1985) are represented by triangles "a" and "b", while those occurring to 20-year-olds correspond to triangles "c" and "d". To estimate the population aged 20 as of June 1, 1985, the event data for ages 19 and 20 must be "separated" in order to match events (represented by the parallelogram produced by concatenating triangles "b" and "c") to the birth cohort that experienced them.

The separation of event data is accomplished through the application of a separation factor (f), and its complement (1 − f). The values of the separation factor, and the particular cases to which they apply, are discussed under the appropriate sections in this chapter.

* This chapter is based essentially on earlier work by Normand Thibault, and on more recent contributions by Y. Edward Shin and Pierre Parent.

Chart 2.1

Transition from a Distribution of Demographic Events by Age and Year to a Distribution by Age and Birth Cohort



Population by Age

The equations for estimating annual population by single years of age, by the cohort component method, can be stated as follows:

For Age 0

$$\hat{P}_{0,t+1} = B_{(t,t+1)} - f_0 D_{0,(t,t+1)} + I_{0,(t,t+1)} - E_{0,(t,t+1)} + N_{0,(t,t+1)} \tag{1}$$

For Ages 1 through 89

$$\hat{P}_{a+1,t+1} = \hat{P}_{a,t} - [(1 - f_a)D_{a,(t,t+1)} + (f_{a+1})D_{a+1,(t,t+1)}] + I_{a,(t,t+1)} - E_{a,(t,t+1)} + N_{a,(t,t+1)} \tag{2}$$

For Ages 90+

$$\hat{P}_{90+,t+1} = \hat{P}_{89+,t} - [(1 - f_{89})D_{89,(t,t+1)} + D_{90+,(t,t+1)}] + I_{89+,(t,t+1)} - E_{89+,(t,t+1)} + N_{89+,(t,t+1)} \tag{3}$$

where: $\hat{P}$ = the estimated population;

B = number of births;

D = number of deaths;

I = number of immigrants;

E = number of emigrants;

N = net interprovincial migration;

f = the separation factor;

t, t+1 = the period June 1, year t to May 31, year t + 1

a = age;

Note: age (a) refers to the age at the time of the event over the period t, t + 1 in the case for deaths; and to the cohort age for immigrants, emigrants and for net interprovincial migration; hence no separation factor (f) is required in the latter cases. Age 0, for immigrants, emigrants and net interprovincial migrants in this equation refers to age at the end of the period, rather than to age at the beginning of the period.

As noted earlier, demographic events recorded by age at the time of event are not suitable for direct use. Moreover, each event requires unique treatment, according to the nature of the data used to generate the estimate. The following sections discuss the manner in which the population components are estimated by age.

(1) Deaths

Estimation of cohort-specific deaths requires the use of the separation factor (f), which varies by age,[1] as indicated in Table 2.1.

TABLE 2.1. Separation Factors for Estimating Cohort-Specific Deaths

Age	f	(1 — f)
0	0.89	0.11
1	0.60	0.40
2 and over	0.50	0.50

The separation factor for age 0 (f = 0.89) indicates that, in any given year, 89 percent of infant deaths occur to infants who were born in the estimation year. The remaining 11 percent of deaths (1 – f) occur to infants who were born in the previous year. The reason for this imbalance has to do with the fact that infant deaths are heavily concentrated in early infancy, and do not, therefore, satisfy the assumption of a rectangular distribution across the age interval. This skewness of the age distribution of deaths extends into the next group, those aged 1-2, but is attenuated to a 60/40 split. By age 2, the assumption of rectangularity is met.

(2) Immigrants and Emigrants

Separation factors are not required for immigration since the data are collated at the outset by month and year of birth, as well as exact month and year of arrival. The ages of the immigrants at June 1 of any year can, therefore, be readily calculated.

Distributions of emigrants by broad age groups (0-17, 18-24, 25-44, 45-64 and 65 and over) are derived from tax files. These age groups are then broken down into five-year age groups on the basis of the age profile of emigrants to the United States. Finally, Sprague's multipliers[2] are applied to generate a distribution by single years of age, which is applied to the total number of emigrants estimated using the method described in Chapter V.

---

[1] The factors for deaths at ages 0 and 1 have been calculated from special tabulations of infant deaths by month of birth, and month and year of death, for the year ending May 31, 1973. It is unlikely that their magnitude has changed to any significant degree since then.

[2] Sprague's multipliers are interpolation coefficients used to subdivide data into halves, fifths or tenths. The first two sets of multipliers are those used in these estimates. These coefficients can be found in H. Shryock, J. Siegel and Associates, *The Methods and Materials of Demography,* U.S. Department of Commerce, Bureau of the Census, Second Printing (rev.), U.S. Government Printing Office, Washington, D.C., 1973, p. 876.

(3) Interprovincial Migrants

To estimate the number of interprovincial migrants by age, the following steps are applied:

(1) Cohort migration ratios, by sex and five-year age group, are calculated using the number of migrants for the most recent intercensal period, by province of origin and destination, using the mobility question from the most recent census ("Where did you live 5 years ago?").

(2) The cohort ratios (calculated in Step 1) are then split into their corresponding age-specific components using Sprague's multipliers. Subsequent regrouping yields migration ratios by age group at the time of the event occurrence.

(3) The number of migrants by age group is calculated by multiplying the migration ratio by the corresponding five-year age group of the most recent Census population.

(4) Based on the number of migrants from the preceding step, a distribution by five-year age group per 100,000 is calculated for each sex, and a single year of age distribution is derived by applying Sprague's multipliers to the grouped data.

(5) A percentage distribution by single years of age is calculated for each broad age-sex group (0-17, 18-24, 25-44, 45-64 and 65 +), using the data obtained in step 4.

(6) The latter distribution, which will remain fixed for the coming 5 years, is applied to the annual migration estimates by broad age group and sex as derived from the taxation file, in order to disaggregate them into a distribution by single years of age.

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

(8) One final adjustment is made to ensure that the number of migrants obtained in step 7, aggregated to broad age groups, is identical to the numbers derived from taxation data, especially in the 0-17 and 18-24 age groups. This process involves eliminating any discrepancy by distributing the difference equally to each single year of age within these two broad age groups.

In principle, steps 1 to 5 are required for each set of origin-destination provinces. However, since there are similarities between some provinces, currently only 34 sets of distributions, instead of some 132 (representing an origin/destination matrix of 12 provinces and territories, minus the redundant cells in the diagonal) that would otherwise be required, are calculated.

Population by Marital Status

The estimation of population by marital status, sex and age is carried out separately for single ($\hat{S}$), married ($\hat{M}$), divorced ($\hat{V}$) and widowed ($\hat{W}$) persons using the following equations:

Single:

$$\begin{aligned}\hat{S}_{a+1,\,t+1} = \hat{S}_{a,\,t} - {} & \tfrac{1}{2}\left[D^{s}_{a,\,(t,\,t+1)} + D^{s}_{a+1,\,(t,\,t+1)}\right] - \\ & \tfrac{1}{2}\left[M^{s}_{a,\,(t,\,t+1)} + M^{s}_{a+1,\,(t,\,t+1)}\right] + \\ & \left[I^{s}_{a,\,(t,\,t+1)} - E^{s}_{a,\,(t,\,t+1)} + N^{s}_{a,\,(t,\,t+1)}\right]\end{aligned} \tag{4}$$

Married:

$$\begin{aligned}\hat{M}_{a+1,\,t+1} = \hat{M}_{a,\,t} - {} & \tfrac{1}{2}\left[D^{m}_{a,\,(t,\,t+1)} + D^{m}_{a+1,\,(t,\,t+1)}\right] + \\ & \tfrac{1}{2}\left[M^{s}_{a,\,(t,\,t+1)} + M^{s}_{a+1,\,(t,\,t+1)}\right] + \\ & \tfrac{1}{2}\left[M^{v}_{a,\,(t,\,t+1)} + M^{v}_{a+1,\,(t,\,t+1)}\right] + \\ & \tfrac{1}{2}\left[M^{w}_{a,\,(t,\,t+1)} + M^{w}_{a+1,\,(t,\,t+1)}\right] - \\ & \tfrac{1}{2}\left[V_{a,\,(t,\,t+1)} + V_{a+1,\,(t,\,t+1)}\right] + \\ & \left[I^{m}_{a,\,(t,\,t+1)} - E^{m}_{a,\,(t,\,t+1)} + N^{m}_{a,\,(t,\,t+1)} - W_{a,\,(t,\,t+1)}\right]\end{aligned} \tag{5}$$

Divorced:

$$\begin{aligned}\hat{V}_{a+1,\,t+1} = \hat{V}_{a,\,t} - {} & \tfrac{1}{2}\left[D^{v}_{a,\,(t,\,t+1)} + D^{v}_{a+1,\,(t,\,t+1)}\right] - \\ & \tfrac{1}{2}\left[M^{v}_{a,\,(t,\,t+1)} + M^{v}_{a+1,\,(t,\,t+1)}\right] + \\ & \tfrac{1}{2}\left[V_{a,\,(t,\,t+1)} + V_{a+1,\,(t,\,t+1)}\right] + \\ & \left[I^{v}_{a,\,(t,\,t+1)} - E^{v}_{a,\,(t,\,t+1)} + N^{v}_{a,\,(t,\,t+1)}\right]\end{aligned} \tag{6}$$

Widowed:

$$\begin{aligned}\hat{W}_{a+1,\,t+1} = \hat{W}_{a,\,t} - {} & \tfrac{1}{2}\left[D^{w}_{a,\,(t,\,t+1)} + D^{w}_{a+1,\,(t,\,t+1)}\right] - \\ & \tfrac{1}{2}\left[M^{w}_{a,\,(t,\,t+1)} + M^{w}_{a+1,\,(t,\,t+1)}\right] + \\ & \left[I^{w}_{a,\,(t,\,t+1)} - E^{w}_{a,\,(t,\,t+1)} + N^{w}_{a,\,(t,\,t+1)} + W_{a,\,(t,\,t+1)}\right]\end{aligned} \tag{7}$$

where:

$\hat{S}$, $\hat{M}$, $\hat{V}$, $\hat{W}$	=	population of marital statuses: single, married, divorced, widowed, respectively, of age a or a + 1 at time t or t + 1;
M, V, W	=	number of marriages, divorces, new widowhoods, respectively, at age a or a + 1 during time interval t, t + 1;

Superscripts s, m, v and w represent "Single", "Married", "Divorced" and "Widowed", respectively. For the meaning of other notations, see formulae used to estimate population by age.

Note: no separation factor was required for emigrants, immigrants, net interprovincial migration and new widowhoods, as the cohort age for these categories was readily available or estimated.

There are two exceptions to the estimation of marital status by age. First, all persons in the age group 0-14 are considered single. Second, for those in the age group 90+ the events (death, widowhood, etc.) are separated into two age groups — those occurring to persons aged 89, and those occurring to persons aged 90 and over. The separation factor ($f = ½$) is applied only to the events in the former age since; for those aged 90 and over, f is equal to one by definition.

A special treatment of marital status by age is required for the following categories:

(1) Immigrants and Emigrants

The marital profile of immigrants by age is available from the records on landed immigrants. The emigrant profile is assumed to be the same as that for immigrants.

(2) Interprovincial Migrants

For each province, the distributions of out-migrants by marital status for a given sex and age are derived from the most recent census data on mobility. These distributions are then applied to the number of out-migrants by province of destination, age and sex estimated in the previous section, to obtain estimates of migrants by marital status for the postcensal years.

(3) Deaths and Marriages

Information on the marital status of deceased persons, and on the previous marital status of newly married persons, is available from Vital Statistics.

(4) New Widowhoods

The term "new widowhoods" denotes a change in marital status from "married" to "widowed", occurring as a result of the death of a spouse in the period t, t + 1. Since the age of the surviving spouse is not recorded, the incidence of "newly widowed" persons is estimated from data on deaths occurring to married persons. Using the age distribution of deaths to married males and married females, compiled by age group (age at the beginning of the estimation period), the surviving spouse is assigned to a five-year age group. Assignment is based on the distribution of husband-wife families by age group of husband and age group of wife from the last census (see Table 2.2).

The numbers of new widows(ers) by age group are estimated as follows:

$$W_{a,\,a+4} = \sum_{a=15}^{90+} D^{m}_{a,\,a+4} \cdot \Phi_{a,\,a+4} \qquad (8)$$

with

$$\Phi_{a,\,a+4} = \frac{F_{a,\,a+4}}{\sum\limits_{a=15}^{90+} F_{a,\,a+4}} \qquad (9)$$

where:

$W_{a,\,a+4}$ = estimated number of new widows(ers) in age group a, a + 4;

$D^{m}_{a,\,a+4}$ = number of deaths of married persons (male or female) in age group a, a + 4;

$\Phi_{a,\,a+4}$ = proportion of husband-wife families by five-year age group of husbands (wives) cross-classified by five-year age group of wives (husbands) a, a + 4;

$F_{a,\,a+4}$ = number of husband-wife families by five-year age group of husbands (wives) cross-classified by five-year age group of wives (husbands) a, a + 4.

An illustration of the above formulae is provided in Table 2.3. This table presents a sample calculation of the incidence of widowhood among females in the 40-44-year age group for the province of Quebec in 1981-1982. The percent distribution of husbands with a 40-44-year old wife is taken from Table 2.2. This is applied to the 1981-82 distribution of male deaths, and the results summed to estimate the number of new widows 40-44-years old. To determine the number of widows in another age group, the appropriate distribution from Table 2.2 is substituted for column 2 of Table 2.3, and the same procedure followed.

TABLE 2.2. Percent Distribution of Husband-Wife Families by Age of Wife and Age of Husband: Quebec, 1981 Census

Age of Husband	Age of Wife								
	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59
15-19	59.1	33.4	4.3	1.4	0.6	0.2	0.2	0.1	0.1
20-24	12.6	73.8	11.7	1.4	0.3	0.1	0.0	0.1	0.0
25-29	1.4	35.0	54.3	8.0	1.0	0.2	0.1	0.0	0.0
30-34	0.3	6.5	37.0	47.7	7.4	0.9	0.2	0.0	0.0
35-39	0.1	1.5	8.6	39.1	43.2	6.4	0.9	0.2	0.0
40-44	0.0	0.4	2.3	10.5	40.3	38.5	6.6	1.1	0.2
45-49	0.0	0.1	0.6	2.7	11.2	37.1	38.6	8.0	1.3
50-54	0.0	0.1	0.2	0.7	2.9	11.1	36.5	39.0	7.9
55-59	0.0	0.0	0.1	0.3	0.8	2.9	11.4	37.2	38.1
60-64	0.0	0.0	0.1	0.1	0.3	0.8	3.0	14.1	39.3
65-69	0.0	0.0	0.0	0.1	0.1	0.3	1.0	4.1	16.9
70-74	0.0	0.0	0.0	0.1	0.1	0.2	0.4	1.4	5.9
75-79	0.0	0.0	0.1	0.1	0.0	0.1	0.2	0.7	2.3
80-84	0.0	0.0	0.1	0.0	0.1	0.0	0.1	0.4	1.3
85-89	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.1	0.6
90+	0.0	0.1	0.3	0.0	0.1	0.0	0.3	0.3	0.5

Age of Husband	60-64	65-69	70-74	75-79	80-84	85-89	90+	**Total**
15-19	0.1	0.2	0.2	0.1	0.0	0.0	0.0	**100.0**
20-24	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**100.0**
25-29	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**100.0**
30-34	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**100.0**
35-39	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**100.0**
40-44	0.1	0.0	0.0	0.0	0.0	0.0	0.0	**100.0**
45-49	0.3	0.1	0.0	0.0	0.0	0.0	0.0	**100.0**
50-54	1.2	0.3	0.1	0.0	0.0	0.0	0.0	**100.0**
55-59	7.7	1.3	0.2	0.0	0.0	0.0	0.0	**100.0**
60-64	33.4	7.6	1.1	0.2	0.0	0.0	0.0	**100.0**
65-69	37.8	31.9	6.7	1.0	0.1	0.0	0.0	**100.0**
70-74	17.9	38.1	29.7	5.5	0.7	0.0	0.0	**100.0**
75-79	6.8	20.3	38.7	26.5	3.9	0.3	0.0	**100.0**
80-84	3.2	8.9	22.1	39.2	21.7	2.7	0.2	**100.0**
85-89	2.1	5.5	13.0	26.2	36.0	15.2	1.1	**100.0**
90+	1.5	4.3	8.7	19.4	28.0	25.7	10.8	**100.0**

Note: The percentages in this table are rounded so as to sum to exactly 100.0.
Source: 1981 Census, unpublished data.

TABLE 2.3. Calculation of the Incidence of Widowhood: An Example for Females Aged 40-44 Years: Quebec, 1981-82

Age of Husband	Deaths of Married Males (1)	Percent Distribution of Husbands with Wife Aged 40-44 Years (2)	Estimates of New Widows Aged 40-44 Years (3) = [(1) × (2)] ÷ 100
15-19	1	0.2	0.002
20-24	38	0.1	0.038
25-29	136	0.2	0.272
30-34	212	0.9	1.908
35-39	300	6.4	19.200
40-44	365	38.5	140.525
45-49	632	37.1	234.472
50-54	1,050	11.1	116.550
55-59	1,564	2.9	45.356
60-64	1,989	0.8	15.912
65-69	2,500	0.3	7.500
70-74	2,521	0.2	5.042
75-79	1,999	0.1	1.999
80-84	1,224	0.0	0
85-89	635	0.0	0
90+	189	0.0	0
TOTAL	**15,355**	-	**588.776** ≈ $\boxed{589}$

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

Evaluation of Estimates

The difference between the census counts and postcensal estimates (referred to as "error of closure" and calculated as in formula 2 of Chapter I) is the criterion used to evaluate the accuracy of estimates by age and marital status.

Estimates by Age and Sex

The errors of closure for the postcensal estimates of Canada's population by age group and sex are presented in Table 2.4. The errors are generally small (around 1 %), and only among the male 20-24-year age group is the error appreciable (3.12 %). This age group, compared to age group 15-19, is characterized by a high census undercoverage rate. An adjustment for undercoverage in the 1976 and 1981 Censuses would have reduced the error of closure for the male 20-24 age group from 3.12 % to – 1.19 %. For other age groups, however, adjustment for undercoverage would not necessarily have reduced the error of closure, since error inherent in the measurement or estimation of other components, accounts for the majority of the total error in these cases.

Estimates by Marital Status

The errors of closure for the adult population (i.e. 15 years of age and older), classified by sex and marital status, are shown in Table 2.5. Married males were underestimated by 3.17 %, while those for single, widowed and divorced males were overestimated by 4.06 %, 1.99 % and 27.09 %, respectively. The errors with respect to the female population are, in most cases, of similar magnitude and direction to those for males.

TABLE 2.4. Error of Closure [1] by Sex and Age Group: Canada, June 1, 1981

Age Group	Male		Female	
	Number	Percent	Number	Percent
0- 4 years	14,289	1.56	13,641	1.57
5- 9 "	– 14,984	– 1.64	– 13,498	– 1.56
10-14 "	– 10,390	– 1.05	– 8,060	– 0.86
15-19 "	– 8,386	– 0.71	– 10,721	– 0.95
20-24 "	36,660	3.12	6,477	0.55
25-29 "	3,094	0.29	2,604	0.24
30-34 "	– 11,721	– 1.15	– 12,428	– 1.22
35-39 "	– 2,888	– 0.35	– 1,975	– 0.24
40-44 "	– 9,570	– 1.42	– 8,036	– 1.21
45-49 "	– 3,432	– 0.54	490	0.08
50-54 "	– 9,057	– 1.46	– 3,355	– 0.54
55-59 "	204	0.04	5,737	0.94
60-64 "	– 5,239	– 1.13	– 2,448	– 0.47
65-69 "	– 2,485	– 0.64	– 5,581	– 1.23
70-74 "	– 733	– 0.26	– 1,591	– 0.45
75-79 "	– 64	– 0.04	– 763	– 0.30
80-84 "	1,372	1.45	1,747	1.08
85 + "	218	0.34	– 211	– 0.16
All Ages	**– 23,112**	**– 0.19**	**– 37,971**	**– 0.31**
Mean Absolute Error [2]		**0.96**		**0.76**

[1] The error of closure is equal to the June 1, 1981 postcensal estimates minus the 1981 Census count adjusted to June 1.

[2] Mean absolute error is the sum of the absolute values of the percent differences divided by the number of categories.

The underestimation of married persons, and overestimation of single and divorced persons, is due primarily to the difference between the Census and the Vital Statistics definition of "married"; the census definition includes persons living in common-law relationships, while the Vital Statistics definition does not.

The impact of common-law unions on marital status estimates deserves a closer look. In the 1981 Census, persons living in common-law relationships were asked to indicate their relationship to person 1 as "common-law partner", and their marital status as "married". However, as shown in Table 2.6, many respondents who stated they were "common-law partners" reported their marital status as other than "married": 44 % reported themselves as "single", 17 % as "divorced" and 3 % as "widowed". Only 36 % of those living in common-law relationships reported that they were "married" and even their marital status is not certain, since some of them may have stated that they were "married" as a consequence of following the instructions given in the census questionnaire, when, in fact, they may have been legally married at some time, become separated, and subsequently entered into the current common-law relationship.

TABLE 2.5. Error of Closure[1] of Estimates of the Population Aged 15 Years and Over by Sex and Marital Status: Canada, June 1, 1981

Marital Status	Male		Female		Both Sexes	
	Number	Percent	Number	Percent	Number	Percent
All Marital Statuses	**– 12,027**	**– 0.13**	**– 30,054**	**– 0.31**	**– 42,081**	**– 0.22**
Single	117,796	4.06	103,190	4.38	220,986	4.21
Married	– 188,749	– 3.17	– 178,309	– 2.98	– 367,058	– 3.07
Widowed	3,980	1.99	– 2,051	– 0.21	1,929	0.17
Divorced	54,946	27.09	47,116	15.85	102,062	20.41

[1] The error of closure is equal to the June 1, 1981 postcensal estimates minus 1981 Census count adjusted to June 1.

TABLE 2.6. Percent Distribution of Persons Living in Common-Law Relationships by Reported Marital Status and Sex: Canada, 1981

	Both Sexes	Male	Female
All Marital Statuses	**100%**	**100%**	**100%**
Single	44	43	45
Married	36	37	34
Widowed	3	2	4
Divorced	17	18	17

Source: 1981 Census of Canada, unpublished data.

By adjusting the population estimates for the misclassification of marital status of common-law partners, errors of closure have been significantly reduced for every category of marital status other than widows, as illustrated in Table 2.7.

TABLE 2.7. Comparison of Error of Closure Between Adjusted and Unadjusted Population Estimates for Persons Living in Common-Law Relationships by Marital Status: Canada, 1981

Male and Female Adjusted and Unadjusted	Population 15 Years of Age and Over				
	All Statuses	Single	Married	Widowed	Divorced
Male					
Unadjusted	**-12,027**	117,796	-188,749	3,980	54,946
Percent	**-0.13**	4.06	-3.17	1.99	27.09
Adjusted	**-12,027**	19,133	-38,751	-787	8,378
Percent	**-0.13**	0.66	-0.68	-0.39	4.13
Female					
Unadjusted	**-30,054**	103,190	-178,309	-2,051	47,116
Percent	**-0.31**	4.38	-2.98	-0.21	15.85
Adjusted	**-30,054**	7,087	-28,311	-11,242	2,412
Percent	**-0.31**	0.30	-0.47	-1.17	0.81

Source: Based on the 1981 Census of Canada, unpublished data.

CHAPTER III

INTERCENSAL ESTIMATES*

Intercensal estimates are produced following each census in order to reconcile postcensal estimates with the census counts, thus assuring the internal consistency of the estimation system. This retrospective revision is also justified since both census counts, and population estimates, are used in the implementation of government revenue sharing programmes. Transfer payments under the Federal-Provincial Fiscal Arrangements and Federal Post-Secondary Education and Health Contribution Programmes are based on census counts in a census year, and postcensal estimates in the postcensal years. Following each census, the intercensal estimates for the two years preceding the census may be used for finalizing interim revenue transfers to provinces.

Estimation Methodology

The following describes the procedures used to generate annual intercensal estimates of total population, of population by age and sex, and of population by age, sex and marital status.

Intercensal Estimates of Total Population

The production of intercensal estimates of total population involves two steps: the calculation of the error of closure as in formula 2 in Chapter I, and; the distribution of this error among the estimates.

The error of closure comes from basically two sources: differences in the amount of undercoverage/overcoverage in successive censuses; and errors in the components of population change over the intercensal period. It is, however, impossible to decompose the error into census and component portions by any empirical means. The assumption is made that the total error of closure is a linear function of the time elapsed since the last census.

Total population estimates for the intercensal years are then obtained by the following formulae:[1]

$$\mathcal{P}_{t-4} = \hat{P}_{t-4} - 1/5\varepsilon \tag{1}$$

$$\mathcal{P}_{t-3} = \hat{P}_{t-3} - 2/5\varepsilon \tag{2}$$

$$\vdots \qquad\qquad \vdots$$

$$\mathcal{P}_{t} = \hat{P}_{t} - 5/5\varepsilon \tag{3}$$

---

* This chapter is based on earlier work by Claude Strohmenger, and on more recent work by Ronald Raby.

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

where:

$\mathcal{P}_{t-4}$ = intercensal population estimate for the year t − 4;

$\hat{P}_{t-4}$ = postcensal population estimate for the year t − 4;

ε = error of closure;

t = the most recent census year;

and t − 4 and t delimit the range between two consecutive censuses.

Intercensal Population Estimates by Age and Sex

The production of intercensal estimates by age and sex involves three steps: the calculation of the error of closure; the distribution of this error, and; the final adjustment of the estimate by sex and age. The error of closure for each sex and single year of age is the difference between the enumerated and estimated population, calculated using the same method as is applied to the total population (see equation 2, Chapter I). Distribution of the error of closure involves additional steps, however, necessitated by the age distribution.

The error of closure for each sex and age cohort is distributed linearly as a function of the time elapsed since the previous census (see Chart 3.1), with the exception of the age groups; "0-4", and "90+". For the age group 0-4 years, the error of closure cannot be distributed over the 5-year period, but only over the number of years elapsed since birth. Accordingly, the error of closure for children aged 4 is distributed over nine half-year periods (4 1/2 years), using 1/9, 3/9, 5/9, 7/9 and 9/9 for the first, second, third, fourth and fifth year from the year of the earlier census. For those aged 3, the error of closure is distributed using 1/7, 3/7, 5/7 and 7/7 over seven 1/2 year periods, or 3 1/2 years, and for those aged 2, using 1/5, 3/5 and 5/5. Finally for 1 year old children, the error of closure is divided using the factors 1/3 and 3/3, covering a 1 1/2 year period.

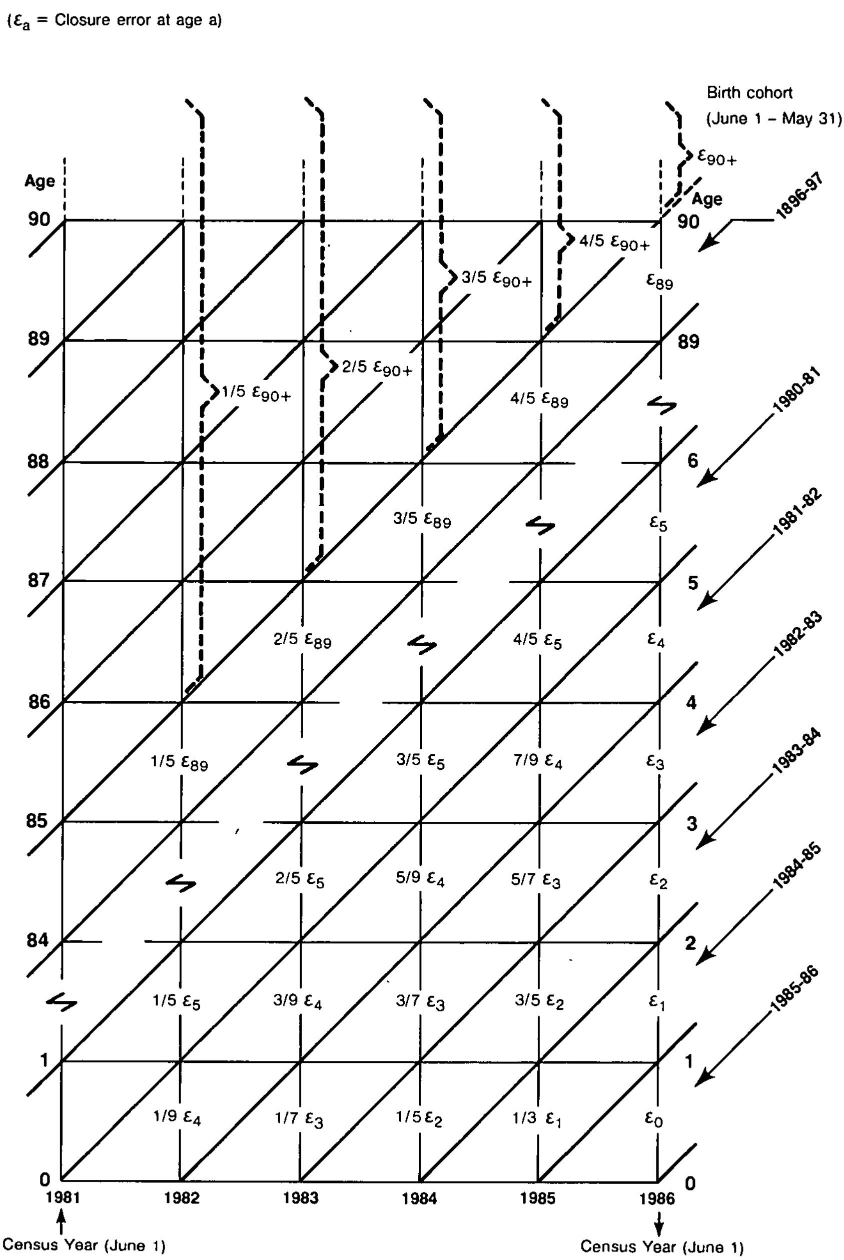
Distribution of the error term for the 90+ age group must take the open-ended nature of the age interval into account. The error of closure for the 90+ age category in 1986 relates to the cohort that was aged 86+ in 1982. Error must be broken down for the period for the single years of age 86 through 89 and the 90+ age group. Error is therefore distributed in proportion to the number of years elapsed since the most recent census (i.e., 1/5, 2/5 ...) and by single years of age prorated over the population in the various age categories (86+ in 1982, 87+ in 1983, ...) as reflected in the final postcensal estimates.

For example, the error of closure (ε age, year) for the year 1983, at age 88, is calculated as:

$$\varepsilon_{88,1983} = \frac{\hat{P}_{88,1983}}{\hat{P}_{87+,1983}} \cdot \frac{2}{5}\,\varepsilon_{90+} \qquad (4)$$

Chart 3.1

Method of Distributing the Error of Closure by Cohort



and for age 90+:

$$\varepsilon_{90+,1983} = \frac{\hat{P}_{90+,1983}}{\hat{P}_{87+,1983}} \cdot \frac{2}{5}\,\varepsilon_{90+} \qquad (5)$$

Having determined the error of closure for each age and sex for the years 1982 through 1986, the intercensal estimates are obtained by subtracting the error from the postcensal estimate for the corresponding year, as in formulae 1 to 3.

When summed, the series of intercensal estimates by age and sex, produced by taking the error of closure into account, differ slightly from the intercensal estimates of total population. These differences are then proportionally distributed among the age cohorts, accounting for the final adjustment made to the estimates.

Intercensal Estimates by Marital Status

The assignment of marital status categories to the intercensal estimates of population by sex and age involves two steps: estimation of the percentage distribution of the population by marital status (single, married, widowed and divorced) for each age by sex, as of June 1, for each intercensal year, and; application of this distribution to the intercensal age-sex specific estimates ($\mathcal{P}$).

The proportional distribution by marital status is calculated from the two last censuses for each sex and age, by province, as follows:

$$\Phi^{ms}_{t-5} = P^{ms}_{t-5} \,/\, P_{t-5} \qquad (6)$$

and

$$\Phi^{ms}_{t} = P^{ms}_{t} / \, P_{t} \qquad (7)$$

where:

ms = marital status: single, married, widowed or divorced;

P_{t-5} and P_t = the number of persons enumerated in the last two censuses;

P^{ms}_{t-5} and P^{ms}_t = the number of persons of marital status ms enumerated in the last two censuses;

Φ^{ms}_{t-5} and Φ^{ms}_t = the proportion of the enumerated population of marital status ms in the 1976 and 1981 Censuses.

The proportional distribution for the intercensal years is estimated by distributing equally the five-year difference associated with a given age-sex category. Thus,

$$\Phi^{ms}_{t-4} = \Phi^{ms}_{t-5} + 1/5\ (\Phi^{ms}_{t} - \Phi^{ms}_{t-5}) \tag{8}$$

$$\Phi^{ms}_{t-3} = \Phi^{ms}_{t-5} + 2/5\ (\Phi^{ms}_{t} - \Phi^{ms}_{t-5}) \tag{9}$$

$$\vdots \qquad\qquad \vdots$$

$$\Phi^{ms}_{t} = \Phi^{ms}_{t-5} + 5/5\ (\Phi^{ms}_{t} - \Phi^{ms}_{t-5}) \tag{10}$$

These proportions are then applied to the intercensal figures for a given age and sex, to obtain estimates by age, sex and marital status.

$$\mathcal{P}^{ms}_{t-4} = \mathcal{P}_{t-4} \cdot \Phi^{ms}_{t-4} \tag{11}$$

where $\mathcal{P}$ = intercensal population estimate by age and sex.

Evaluation of Estimates

Comparison of postcensal estimates and intercensal estimates of total population is made in Table 3.1. As expected, the difference increases with each year because the errors accumulate as a function of the time elapsed from the base census. Their overall magnitude is small, however, attesting to the high accuracy of the annual postcensal estimates, with respect to the census counts.

TABLE 3.1. Percent Differences Between Intercensal and Postcensal Total Population Estimates: Canada, Provinces and Territories, 1977 to 1981

Geographic Area	1977	1978	1979	1980	1981
Canada	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**
Newfoundland	−0.25	−0.51	−0.76	−1.00	−1.25
Prince Edward Island	0.06	0.13	0.19	0.25	0.31
Nova Scotia	0.01	0.01	0.02	0.02	0.03
New Brunswick	0.06	0.11	0.17	0.22	0.28
Quebec	0.12	0.24	0.36	0.47	0.58
Ontario	0.08	−0.15	−0.22	−0.29	−0.37
Manitoba	0.17	−0.33	−0.50	−0.67	−0.83
Saskatchewan	0.11	0.21	0.32	0.42	0.52
Alberta	0.56	−0.09	1.58	2.01	2.41
British Columbia	0.05	0.10	0.14	0.18	0.22
Yukon	0.45	0.87	1.32	1.75	2.11
Northwest Territories	1.12	2.35	3.49	4.59	5.60

Differences between postcensal and intercensal estimates are larger when considering characteristics by age and by marital status, as indicated in Tables 3.2 and 3.3, respectively. For example, the differences reach .2 % to 3 % among young males 20-24 years old, precisely the ones who also record the highest census undercoverage rate. Among divorced males, differences range from 9 % to 27 %. Though not shown here, the differences between the two estimates are very large, 48 % and 16 %, among those married in the age groups 15-19 and 20-24, respectively. The chief distorting factor seems to be common-law unions. While such unions are included as married in the census counts, there is no provision for common-law unions in the marriage registration records and, hence, they are not accounted for in the postcensal estimates of population by marital status.

TABLE 3.2. Percent Differences Between Intercensal and Postcensal Population Estimates by 5-Year Age Groups and Sex: Canada, 1977 to 1981

Age Groups	1977	1978	1979	1980	1981
			Males		
0- 4	0.22	0.22	-0.07	-0.62	-1.56
5- 9	0.30	0.60	0.85	1.21	1.64
10-14	0.18	0.54	0.82	0.89	1.05
15-19	-0.63	-1.09	-0.97	-0.38	0.71
20-24	-0.17	-0.57	-1.40	-2.21	-3.12
25-29	0.23	0.43	0.62	0.37	-0.29
30-34	0.06	0.26	0.59	0.93	1.15
35-39	0.27	0.42	0.39	0.27	0.35
40-44	0.08	0.35	0.82	1.09	1.42
45-49	0.30	0.48	0.42	0.52	0.54
50-54	0.06	0.35	0.65	0.95	1.46
55-59	0.11	0.05	0.15	0.12	-0.04
60-64	0.23	0.45	0.55	0.82	1.13
65-69	-0.09	-0.04	0.09	0.07	0.64
70-74	0.02	0.07	0.21	0.48	0.26
75-79	-0.18	-0.42	-0.48	-0.32	0.04
80-84	0.19	0.44	0.29	-0.47	-1.45
85-89	-0.31	-0.40	-0.06	1.06	2.34
90+	-0.51	-1.26	-2.33	-3.95	-6.37

TABLE 3.2. Percent Differences Between Intercensal and Postcensal Population Estimates by 5-Year Age Groups and Sex: Canada, 1977 to 1981 - Concluded

Age Groups	1977	1978	1979	1980	1981
			Females		
0- 4	0.19	0.14	−0.17	−0.69	−1.57
5- 9	0.25	0.50	0.71	1.12	1.56
10-14	0.13	0.36	0.61	0.66	0.86
15-19	−0.15	−0.29	−0.17	0.29	0.95
20-24	−0.10	−0.20	−0.44	−0.53	−0.55
25-29	0.22	0.37	0.51	0.33	−0.24
30-34	0.01	0.11	0.42	0.86	1.22
35-39	0.23	0.38	0.34	0.24	0.24
40-44	−0.07	0.08	0.48	0.77	1.21
45-49	0.12	0.09	−0.05	−0.05	−0.08
50-54	−0.17	−0.10	−0.09	0.12	0.54
55-59	−0.04	−0.28	−0.36	−0.60	−0.94
60-64	0.32	0.46	0.46	0.48	0.47
65-69	−0.06	0.04	0.25	0.45	1.23
70-74	0.05	0.19	0.30	0.73	0.45
75-79	−0.24	−0.62	−0.62	−0.40	0.30
80-84	0.21	0.59	0.47	−0.25	−1.08
85-89	−0.06	−0.32	0.05	1.19	2.08
90+	−0.17	−0.89	−1.54	−2.45	−3.69

TABLE 3.3. Percent Differences Between Intercensal and Postcensal Estimates of Population Aged 15 Years and Over by Marital Status and Sex: Canada, 1977 to 1981

Marital Status and Sex	1977	1978	1979	1980	1981
Both Sexes					
Single	0.62	1.52	2.53	3.29	4.21
Married	-0.49	-1.11	-1.79	-2.39	-3.07
Widowed	0.11	0.35	0.30	0.30	0.17
Divorced	6.64	11.69	15.55	18.07	20.41
Male					
Single	0.61	1.47	2.41	3.17	4.06
Married	-0.50	-1.13	-1.86	-2.47	-3.17
Widowed	0.13	0.63	1.08	1.57	1.99
Divorced	8.88	15.55	20.60	24.02	27.09
Female					
Single	0.64	1.59	2.67	3.44	4.38
Married	-0.48	-1.08	-1.73	-2.32	-2.98
Widowed	0.11	0.29	0.14	0.03	-0.21
Divorced	5.17	9.12	12.15	14.04	15.85

CHAPTER IV

INTERNAL MIGRATION*

Canada does not have a system for the registration of migrants comparable to that for the registration of births, deaths and other vital events. The main source of information on migration is the census question on the respondent's residency five years prior to the census date. As presently worded, however, this question allows for the recording of only one move over the five-year interval. Therefore, to the extent that some migrants may move more than once during this period, these data clearly understate the incidence of annual migration. Neither do the data capture those migrants who move away from, and later return to, their province of origin within the five-year interval, as well as with respect to migrants who either emigrate outside Canada, or who die after an interprovincial move. For purposes of preparing quarterly and annual population estimates, therefore, this five-year migration information is not very helpful. Consequently, to determine annual and quarterly migration flows, Statistics Canada makes use of information available from administrative records.

From 1956 to 1976 the required information was provided by family allowance data.[1] Since 1976, this has been supplemented with data from personal income tax files. Today there are two sets of migration estimates: preliminary estimates based on data from family allowance files, and final estimates based on data from taxation files.

The main advantages of the family allowance data are their high quality in terms of coverage and accuracy, as well as their timely availability. As such, they are particularly suited to the requirements for the production of quarterly and annual *preliminary* estimates for Canada, the provinces and territories. By comparison, the taxation files lag the family allowance files by about 12 months, but provide additional pertinent information (especially on adult migrants and by age and sex). More importantly, they provide migration data not only by provinces, but also for smaller geographical units such as census divisions and census metropolitan areas. In order to ensure consistency of the migration estimates at various levels of geographical disaggregation, it was deemed preferable to use taxation data for *final* migration estimates.[2]

---

* This chapter is a revised version of a paper presented by Anatole Romaniuc, Ronald Raby and Pierre Parent, at the annual meeting of the Population Association of America, held at Minneapolis, Minnesota, May 1984.

[1] Y. Kasahara, *The Flow of Migration Among the Provinces of Canada — 1951-1961,* Canadian Political Science Association Conference on Statistics, Montreal, 1961.

[2] Note that under the current processing system of family allowances, it is possible to generate, the *number of recipients* of family allowances for subprovincial units, *but not the number of migrants.*

Estimation Methodology and Data Sources

Family Allowance-Based Method

Description of Family Allowance Files

In Canada, every child under the age of 18 years is entitled to a statutory monthly payment known as "Family Allowance", providing the child is wholly or substantially supported by a parent who is either a Canadian citizen, a landed immigrant, or who has been allowed to enter and remain in Canada for at least one year, and who has been subject to Canadian income taxation. Each month, every eligible family is sent a family allowance cheque by mail. To continue receiving their family allowance cheques, recipients must notify the regional office of Health and Welfare Canada of any changes of address. These notifications of change of address form the basis of the family allowance-based migration recording system. The data can be compiled by province of origin and destination, in terms of the number of families by size (the number of children per family receiving the allowance), and further into the month of the event.

Of the files produced from family allowance records, four are used in migration estimates: child population files, M0023 and the monthly reports; and two corresponding child migration files, M0024 and M0013.

(1) M0024 Child Migration File

This file shows the month in which the migration actually took place. Although it uses the same change of address information as M0013, it is compiled somewhat differently, in that corrections to the real month of migration, as shown on the change of address notice, can be made for up to 24 months after the move has taken place. Corrected monthly interprovincial migration statistics covering a 24-month period are compiled every June and December. The data for the most recent six-month period (referred to as Edition 1) are preliminary, since they are presented for the first time. Each of the three remaining releases (namely, editions 2,3 and 4) incorporate revisions made at six-month intervals to previously released data. Edition 4 is final.

M0024 data are released about three months after the end of the reference periods (Jan. — June, and July — Dec.). Hence, the preliminary (Edition 1) M0024 data for a specific month are not available until 3 to 8 months after a move takes place, while final monthly data (Edition 4) are not available until 21 to 26 months after the move. The data are compiled by province of departure (referred to as outward direction) and by province of entry (inward direction).

Because of its later production, M0024 is not used for preliminary estimates of migration. Rather, its primary use is for disaggregating annual migration data from tax files into quarterly data. It also serves as a reference base for evaluating M0013 data, in addition to providing some useful information not available from M0013. The additional data includes the distribution of child-migrants by single years of age, and child migration flows into and out of Canada.

(2) M0013 Child Migration File

The M0013 file is forwarded to Statistics Canada at the end of each month. It is based on the month in which the change of address is processed by Health and Welfare Canada (there is approximately a two-month administrative action delay between the date of the actual move and issuance of payment to the new address). Because of its timeliness, this file is used in the preparation of the *preliminary* estimates.

(3) M0023 Child Population File, and the Monthly Report of Family Allowance Recipients.

In addition to information on changes of address, Health and Welfare provides Statistics Canada with a count of children *receiving* the family allowance. The monthly report (formerly known as the M0011 file) is received together with the M0013 file. The updates, identified as M0023, show the number of children *entitled* to family allowances, and are received later with the M0024 file. These data on the child population at risk of migration are precisely the data required to calculate the rates of child migration.

Coverage and Consistency of the Files

The quality of the family allowance data base is contingent upon two factors: the comprehensiveness of coverage of the child population under the age of 18; and the internal consistency of the different child migration files (how the various family allowance files compare to each other).

The comprehensiveness of coverage can be judged by comparing the family-allowance-based numbers of children under age 18, with those of census counts as of June 1, 1981. As shown in Table 4.1, the two are almost identical, suggesting that the family allowance coverage is comparable to that of the census, particularly for those who are under the age of 15. For those aged 15 to 17 years, the family allowance data coverage for Canada is 97 %, and, excluding the territories, the highest and lowest rates of coverage among the provinces are for Québec and Alberta, at 98.2 % and 93.6 %, respectively. This undercoverage is explained by the fact that children not wholly or substantially supported by a parent are not entitled to the family allowance, and are, therefore, not included in the family allowance files.

Table 4.2 compares the consistency of the various migration files generated from family allowance data, and shows that the differences are very small.

TABLE 4.1. Ratio Between Family Allowance Files [1] and Census Counts [2], for the Child Population by Broad Age Groups: Canada, Provinces and Territories, June 1, 1981

Geographic Area	Age 0-17		Age 0-14	Age 15-17
	Monthly Report	M0023	M0023	M0023
Canada	**0.993**	**1.000**	**1.008**	**0.970**
Newfoundland	0.999	1.003	1.012	0.964
Prince Edward Island	0.996	0.998	1.004	0.974
Nova Scotia	0.997	1.001	1.009	0.971
New Brunswick	1.001	1.003	1.010	0.975
Quebec	0.996	1.002	1.007	0.982
Ontario	0.992	0.999	1.007	0.970
Manitoba	0.999	1.005	1.013	0.971
Saskatchewan	0.993	0.998	1.006	0.963
Alberta	0.977	0.989	1.001	0.936
British Columbia	0.999	1.007	1.017	0.968
Yukon	0.991	1.012	1.030	0.929
Northwest Territories	1.016	1.033	1.044	0.979

[1] Two different family allowance files are used.
Monthly Report: Gives the number of children **receiving** the family allowance in June, 1981.
M0023:Gives the number of children **entitled** to the family allowance as of June 1, 1981.
[2] The Census was held on June 3. Census data used here are adjusted to June 1, 1981.

Generally, the number of migrants compiled by province of departure (M0024, outward) falls short of that compiled by province of arrival (M0024, inward),[3] but only by a small margin. By contrast, the M0013 migrant counts, which are compiled only by province of departure, slightly exceed the M0024 province of arrival counts of migrants in most of the provinces. The discrepancy between these latter two files is narrowed when the M0013 migrant counts are lagged by two months to correspond more closely to the month of migration as reported in the M0024.

As the foregoing reveals, there is very close agreement between the two migration files, and the two child population files. Unless there is a built-in bias in the system, this suggests that the family allowance files provide a reliable

[3] One would expect the M0024 inward file to be more accurate given the various safeguards usually taken in rerouting payments to the new addresses. For example, before opening a new account for an in-migrant to a province, the Health and Welfare office in the province of arrival ensures that the beneficiary's account in the province of departure has been closed. Also, a beneficiary leaving a province may or may not inform the Health and Welfare office of that province of his/her departure, but he/she most likely would report to the Health and Welfare office in the province of arrival to ensure receipt of payment. It might also be noted that even if Health and Welfare were notified of a recipient's intended departure, the recipient might decide either not to migrate, or to move to a province other than the one originally indicated.

TABLE 4.2. Ratio Between Different Family Allowance Files for the Number of Child Interprovincial Out-Migrants by Province or Territory of Origin, 1980-81, 1981-82 and 1982-83

Geographic Area	M0013[1] / M0024 In[2]	M0013 Lag[3] / M0024 In	M0024 Out[4] / M0024 In
	1980-1981		
Canada	**1.004**	**1.009**	**0.976**
Newfoundland	1.026	1.043	1.008
Prince Edward Island	1.016	1.010	0.963
Nova Scotia	1.000	1.020	0.955
New Brunswick	1.002	0.991	0.986
Quebec	1.021	1.008	0.972
Ontario	1.023	1.020	0.980
Manitoba	1.027	1.003	0.992
Saskatchewan	1.013	0.997	0.975
Alberta	0.970	0.994	0.959
British Columbia	0.973	1.013	0.964
Yukon	1.048	1.000	0.953
Northwest Territories	1.040	1.013	0.986
	1981-1982		
Canada	**1.021**	**1.002**	**0.968**
Newfoundland	1.024	1.008	1.025
Prince Edward Island	0.998	0.972	0.969
Nova Scotia	1.029	0.978	0.986
New Brunswick	1.068	1.018	0.989
Quebec	1.018	1.029	0.982
Ontario	1.044	1.002	0.963
Manitoba	1.034	1.004	0.986
Saskatchewan	1.008	0.985	0.973
Alberta	0.981	0.988	0.933
British Columbia	1.023	1.013	0.970
Yukon	1.034	1.072	1.025
Northwest Territories	1.022	0.983	0.953

TABLE 4.2. Ratio Between Different Family Allowance Files for the Number of Child Interprovincial Out-Migrants by Province or Territory of Origin, 1980-81, 1981-82 and 1982-83 - Concluded

Geographic Area	1982-1983		
	M0013[1] / M0024 In[2]	M0013 Lag[3] / M0024 In	M0024 Out[4] / M0024 In
Canada	**1.045**	**1.019**	**0.980**
Newfoundland	1.050	1.001	0.976
Prince Edward Island	1.074	1.036	0.974
Nova Scotia	1.026	1.004	0.970
New Brunswick	1.041	1.005	0.982
Quebec	1.036	1.009	0.968
Ontario	1.063	1.013	0.986
Manitoba	1.059	1.033	0.988
Saskatchewan	1.044	1.011	0.984
Alberta	1.027	1.040	0.979
British Columbia	1.052	1.007	0.977
Yukon	1.045	1.020	0.994
Northwest Territories	1.011	1.074	1.016

[1] Derived from M0013 file.
[2] Derived from M0024 inward direction file.
[3] Derived from M0013 file, and lagged by two months to correspond more closely to the exact month of migration as reported in the M0024 files.
[4] Derived from M0024 outward direction file.

means of measuring migration. The consistency among the files also indicates that one file could be substituted for any other. For the preliminary estimates, the overriding selection criterion is the timeliness of the file. The M0013 file, available about two months after a move has taken place, has a definite advantage over the M0024 updates.

Estimation of Adult Migration

Family allowance files contain no information on adult migrants. Accordingly, the number of adult migrants must be estimated. This is done as follows:

First, ratios of adult out-migration rates to child out-migration rates are calculated as follows from taxation data for the most recent year, usually 1 or 2 years prior to the estimation period:

$$f_{(j,\,k)} = \frac{M_{(j,\,k),\,18+}}{\hat{P}_{j,\,18+}} \div \frac{M_{(j,\,k),\,0\text{-}17}}{\hat{P}_{j,\,0\text{-}17)}} \qquad (1)$$

where:

$f_{(j,\,k)}$ = estimation factor for adult migrants from province of origin j to province of destination k;

$M_{(j,\,k),\,18+}$ = number of adults (aged 18+), out-migrants from province j to province k; from Revenue Canada income tax file;

$M_{(j,\,k),\,0\text{-}17}$ = number of children (aged 0-17), out-migrants from province j to province k; from Revenue Canada income tax file;

$\hat{P}_{j,\,18+}$ = number of adults in province j; Demography Division estimates;

$\hat{P}_{j,\,0\text{-}17}$ = number of children in province j; Demography Division estimates.

Next, the child out-migration rate by province of destination is calculated for each j^{th} province from family allowance data (the first factor on the right hand side of equation 2).

An estimate of the number of adult out-migrants is calculated by multiplying the estimation factor $f_{(j,k)}$, by the child migration rate, by the estimated population 18+ (equation 2).

$$\hat{M}_{(j,\,k),18+} = \frac{M_{(j,\,k),\,0\text{-}17}}{P_{j,\,0\text{-}17}} \cdot f_{(j,k)} \cdot P_{j,18+} \qquad (2)$$

where:

$\hat{M}_{(j,\,k),\,18+}$ = estimated number of adult (aged 18+) out-migrants from province j to province k;

$M_{(j,\,k),\,0\text{-}17}$ = number of child (aged 0-17) out-migrants from province j to province k, based on the family allowance file;

$P_{j,\,18+}$ = estimated number of adults in province j, computed "residually" from total population estimates (Demography Division), and child population counts (family allowance);

$P_{j,\,0\text{-}17}$ = total number of children receiving family allowance payments in province j; based on family allowance file.

Next, equation 3 is used to derive the total number of out-migrants from province j to province k ($\hat{M}_{(j,k)}$):

$$\hat{M}_{(j,\,k)} = M_{(j,\,k),\,0\text{-}17} + \hat{M}_{(j,\,k),\,18+} \qquad (3)$$

Finally, the total number of out-migrants from j to all other provinces ($\hat{M}_j$) is obtained from equation (4):

$$\hat{M}_j = \sum_{k \neq j} \hat{M}_{(j,\,k)} \tag{4}$$

Conversely, summing over the individual provinces of origin (j), one can calculate the total number of in-migrants to a given province (k). Net migration is then the excess (or deficit) of in-migrants over out-migrants.

In the above method of estimating interprovincial migration, the f factor used to estimate adult migration is of key significance. There are four important issues concerning the use of the f factor. The first is whether f should be calculated for the province of origin only (f_j), or for the provinces of origin and of destination ($f_{(j,k)}$). In earlier years, f was calculated only for the province of origin, but as more comprehensive data became available, it became apparent that the ratios of adult to child out-migration rates varied significantly, not only with the province of origin (f_j), but also with the province of destination ($f_{(j,k)}$). This is illustrated in Chart 4.1 for the province of Quebec. Consequently, the decision was made to calculate the estimation factor f by both province of origin and province of destination ($f_{(j,k)}$).

The second issue is the choice of data base. Before 1981, the only data used were those from the most recent census. As noted earlier, however, the census question on place of residence, perforce, ignores multiple moves over the five-year interval between censuses. Year to year fluctuations are not reflected in the 5-year "aggregates". Moreover, on average, it reflects the migration trends of the previous five years, which may not be comparable to the estimation period. These points are illustrated in Chart 4.1 and Table 4.3. For example, in Alberta, f_j values varied between 1.08 (1981 Census) and 1.24 (1982-83 Taxation Statistics). As annual age-specific data on migrants became available from income tax records, the decision was made to base the calculation of $f_{(j,k)}$ on this data. This annual data shows more recent migration patterns, compared to the census data, which is produced only every fifth year.

The third issue, unresolved so far, relates to the lag of about 12 months between the production of the earliest version of Family Allowance file M0013, and the availability of the tax record data. The use of an f factor based on data from previous years, rather than on current data, is a potential source of error, the extent of which cannot be known in advance. For example, the current-year adult migration can be distorted when the f_j values from the previous year are used, as happened in the case of Alberta during the economically volatile period 1981-82 to 1983-84. The scatter diagram in Chart 4.2, showing f_j values for specified data sources and years, reveals a fairly high correlation coefficient, notwithstanding the fact that in some instances f_j values for individual provinces fluctuate significantly from year to year. A sudden change in the economic climate in a province can result in a rather radical departure from the established pattern of migration for that province.

Chart 4.1

Variation Between Values of $f_{(j,k)}$ for Out-Migrants from Quebec to Other Provinces for Selected Years, Based on Tax Files and the 1976 and 1981 Censuses

Chart 4.1
Variation Between Values of $f_{(j,k)}$ for Out-Migrants from Quebec to Other Provinces for Selected Years, Based on Tax Files and the 1976 and 1981 Censuses

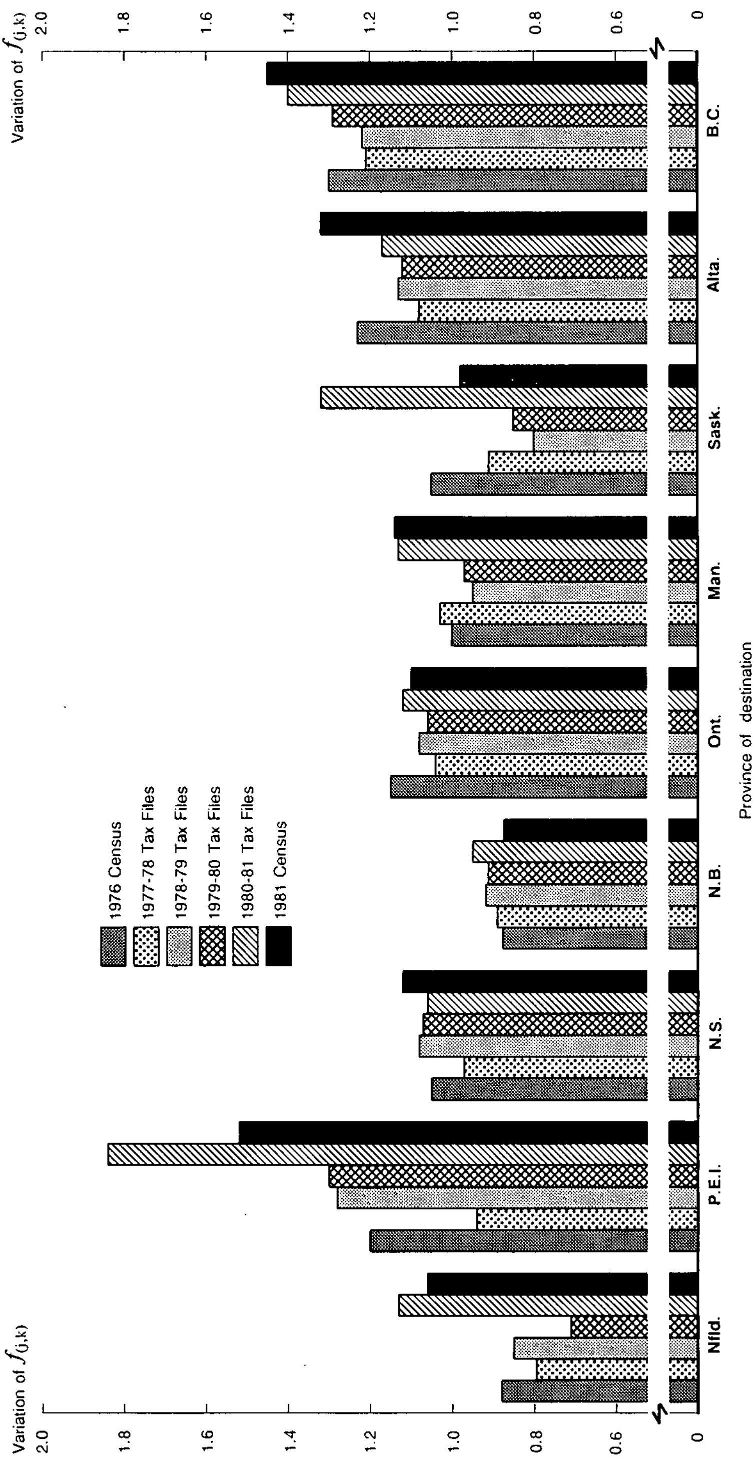


TABLE 4.3. Values of f_j Based on Censuses and on Taxation Records: Provinces and Territories: Selected Years

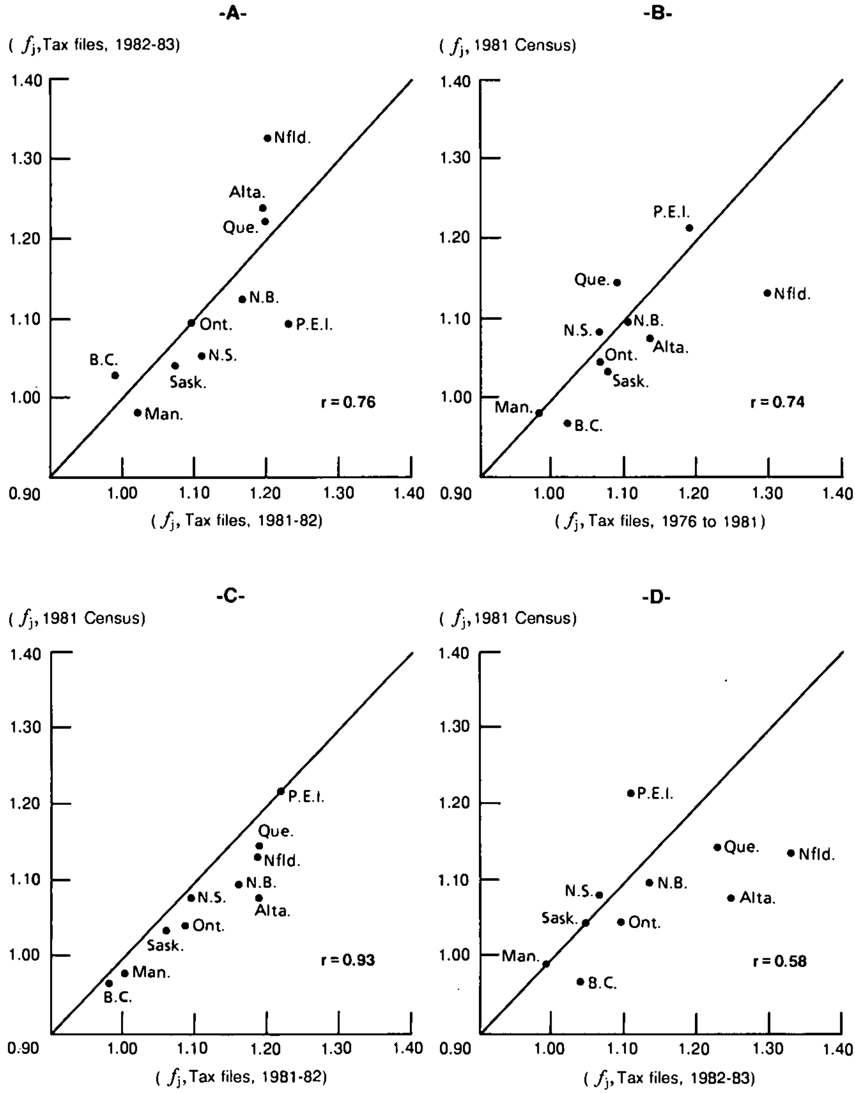
Geographic Area	Census Based f_j	
	1976	1981
Newfoundland	1.2773	1.1372
Prince Edward Island	1.2402	1.2181
Nova Scotia	1.0453	1.0800
New Brunswick	1.0831	1.1010
Quebec	1.1045	1.1448
Ontario	1.0024	1.0460
Manitoba	1.0425	0.9814
Saskatchewan	1.0494	1.0383
Alberta	1.1078	1.0789
British Columbia	0.9166	0.9673
Yukon	1.2108	1.0851
Northwest Territories	1.8753	1.6762

Geographic Area	Income Tax Based f_j			
	1980-81	1981-82	1982-83	1983-84
Newfoundland	1.3648	1.1876	1.3279	1.2638
Prince Edward Island	1.2744	1.2205	1.1053	1.1770
Nova Scotia	1.1230	1.0933	1.0526	1.0453
New Brunswick	1.1645	1.1578	1.1379	1.1294
Quebec	1.1475	1.1880	1.2111	1.1753
Ontario	1.1165	1.0868	1.0950	1.1237
Manitoba	1.0285	1.0060	0.9858	0.9973
Saskatchewan	1.0530	1.0633	1.0878	1.0181
Alberta	1.1813	1.1856	1.2432	1.1542
British Columbia	1.0394	0.9793	1.0297	1.0188
Yukon	1.1889	1.2862	1.1144	1.0094
Northwest Territories	1.5161	1.5516	1.5965	1.4602

Source: Statistics Canada, Demography Division

Chart 4.2

Variation Between Values of f_j by Province for Specified Sources and Years



The fourth and final issue is the question of the reliability of the child migration information derived from tax records, as opposed to the information on adult migration. This remains an important issue, although in the absence of relevant data, it cannot be addressed here.

Income Tax-Based Method

Description of Income Tax Files

All individuals receiving an annual income above a specified minimum are required to file an income tax return by the end of April each year. Each return filed contains information on a number of demographic characteristics, such as year of birth, sex and marital status. Most importantly, from the point of view of determining mobility, each tax return contains the filer's address at the time of filing.[4] Individual tax records for two consecutive years are matched, in order to select those who filed for the two-year period. In recent years, approximately 90 % of the individuals who filed a return in a given year, had also done so in the previous year. This is slightly less than the maximum of 93 %, reached in 1981-82.

Table 4.4 gives some indication of the comprehensiveness of the coverage of "matched" tax filers. For example, 78 % of the population over 18 years of age filed tax returns in both 1981 and 1982, and 73 % in both 1983 and 1984. When the filers' dependents are also taken into consideration, the percentage of population accounted for in each of the two-year periods was 95 % and 88 %, respectively. The coverage varies, however, by age group and sex. Table 4.5, for example, shows that of all returns filed in 1983 and 1984 by individuals over the age of 18 years, a successful match was obtained for 78 % of returns filed by males, but for only 69 % of those filed by females. Thus, though not universal, the comprehensiveness of coverage makes the tax data base a viable source of information on migration.

Estimation of Interprovincial Migrants from Income Tax Data

The estimation of interprovincial migration from income tax data involves three steps: determination of the mobility status of tax filers; imputation of migratory status and demographic characteristics to dependents; and adjustment for coverage. This process is outlined in schematic form in Chart 4.3.

(1) Determination of Mobility Status of Tax Filers

Among the matched tax filers, migrants are identified by comparing the mailing addresses for the two consecutive years. Since the "deadline" for

[4] For a more detailed account of income tax-based migration data, the reader is referred to D.A. Norris and L.D. Standish, *A Technical Report on the Development of Migration Data from Taxation Records*, Administrative Data Development Division, Statistics Canada, Ottawa, May 1983.

TABLE 4.4. Coverage Ratios of Matched Tax Filers, and Matched Filers and Dependents: Canada, Provinces and Territories, 1980-81 to 1983-84

Geographic Area	Tax Filers [1]			
	1980-81	1981-82	1982-83	1983-84
Canada	**0.750**	**0.776**	**0.766**	**0.733**
Newfoundland	0.732	0.762	0.749	0.713
Prince Edward Island	0.704	0.730	0.723	0.702
Nova Scotia	0.723	0.749	0.742	0.707
New Brunswick	0.734	0.762	0.746	0.715
Quebec	0.706	0.736	0.736	0.714
Ontario	0.781	0.801	0.791	0.758
Manitoba	0.828	0.852	0.842	0.804
Saskatchewan	0.726	0.760	0.755	0.736
Alberta	0.746	0.777	0.770	0.728
British Columbia	0.750	0.782	0.744	0.694
Yukon	0.773	0.804	0.811	0.770
Northwest Territories	0.743	0.777	0.786	0.749
	Tax Filers and Dependents [2]			
Canada	**0.951**	**0.945**	**0.923**	**0.883**
Newfoundland	0.959	0.960	0.940	0.895
Prince Edward Island	0.902	0.907	0.894	0.859
Nova Scotia	0.947	0.946	0.930	0.885
New Brunswick	0.945	0.942	0.919	0.878
Quebec	0.914	0.914	0.908	0.882
Ontario	0.985	0.974	0.952	0.909
Manitoba	1.015	1.011	0.979	0.927
Saskatchewan	0.931	0.927	0.903	0.875
Alberta	0.916	0.914	0.890	0.844
British Columbia	0.937	0.935	0.879	0.826
Yukon	0.893	0.902	0.904	0.859
Northwest Territories	0.886	0.893	0.900	0.876

[1] Coverage ratios of matched tax filers are calculated by dividing the number of tax filers matched for two consecutive years by the estimated or enumerated population aged 18 and over (16 and over in 1980-81) at the beginning of the period.
[2] Coverage ratios of matched filers and dependents are calculated by dividing the total number of matched filers and dependents for two consecutive years by the total estimated or enumerated population at the beginning of the period.
Source: Statistics Canada, Small Area and Administrative Data Division.

TABLE 4.5. Coverage Ratios of Matched Tax Filers, Excluding and Including Dependents, by Age Group and Sex: Canada, 1981-82 to 1983-84

Age Group and Sex	Tax Filers [1]		
	1981-82	1982-83	1983-84
Males			
0-17	0.045	0.042	0.027
18-24	0.825	0.789	0.708
25-44	0.885	0.863	0.821
45-64	0.879	0.865	0.836
65+	0.620	0.629	0.627
18+	0.840	0.822	0.781
All ages	**0.600**	**0.600**	**0.572**
Females			
0-17	0.031	0.031	0.021
18-24	0.807	0.791	0.729
25-44	0.858	0.849	0.822
45-64	0.622	0.688	0.616
65+	0.380	0.399	0.402
18+	0.714	0.712	0.687
All ages	**0.568**	**0.532**	**0.515**
	Tax Filers and Dependents [2]		
Males			
0-17	0.940	0.922	0.889
18-24	0.985	0.945	0.872
25-44	0.901	0.879	0.837
45-64	0.899	0.885	0.857
65+	0.636	0.645	0.643
18+	0.885	0.866	0.826
All ages	**0.901**	**0.882**	**0.841**
Females			
0-17	0.942	0.925	0.897
18-24	1.120	1.087	1.024
25-44	1.178	1.137	1.086
45-64	0.949	0.927	0.899
65+	0.508	0.513	0.510
18+	1.006	0.978	0.925
All ages	**0.988**	**0.964**	**0.925**

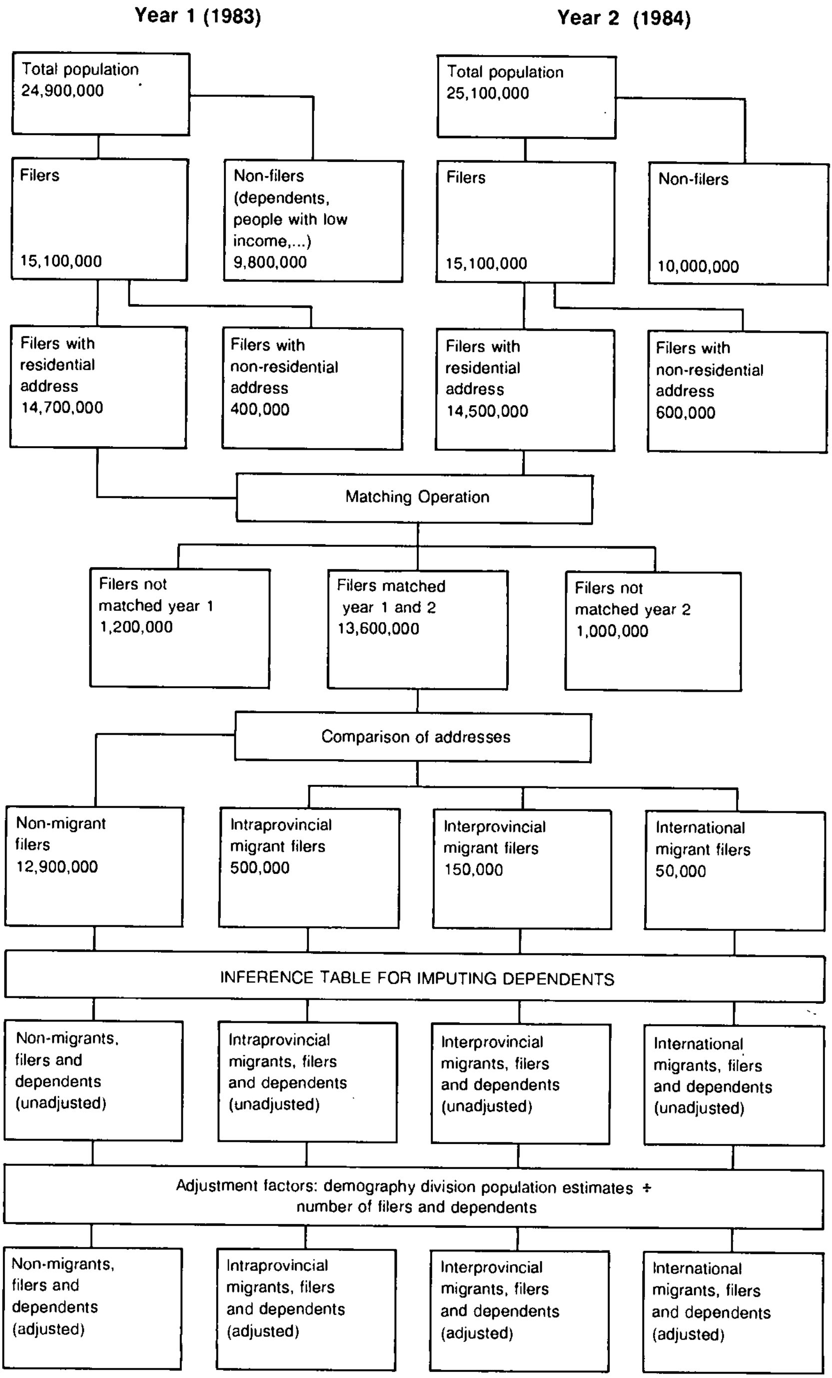
[1] Coverage ratios for each age group and sex are calculated by dividing the number of tax filers matched for two consecutive years by the estimated or enumerated population at the beginning of the period.

[2] Coverage ratios, for each age group and sex, are calculated by dividing the number of matched filers and dependents for two consecutive years by the estimated or enumerated population at the beginning of the period.

Source: Statistics Canada, Small Area and Administrative Data Division.

Chart 4.3

Description of the Process of Estimating Migrants from Tax Files



filing returns is the end of April, the period of migration covers approximately the one-year period from April to April. The migration status of approximately 75 % of the Canadian adult population (18 years of age and over since 1981; 16 years and over prior to 1981) can be determined directly from individual income tax file addresses.

(2) Imputation of the Migratory Status and Characteristics of Dependents

The main source file from which migration information is obtained contains no direct information on tax filers' dependents. However, this information can be imputed from the dollar value of total personal exemptions. Imputation is based on a second Revenue Canada tax record file, representing about 3 % of total tax returns. This sample, containing approximately 350,000 filers, focuses on the specifics pertaining to various personal exemptions, and is stratified by: (a) duration of residence in Canada (i.e., full year, or part thereof); (b) age group (0-17, 18-24, 25-44, 45-64, 65 +); (c) sex; and (d) marital status. Tax filers, resident in Canada for the full year, are classified into approximately 100 exemption classes, of which 50 correspond to various combinations of full exemption (and are, therefore, called exact dollar exemption classes). In fact, total exemptions can be broken down into exact dollar values for 90% of the tax filers; the remaining 50 exemption classes correspond to various combinations, which include partial exemption. In all, there were a total of more than 2,000 cross-classifications. (100 exemption classes x 2 sexes x 5 age groups x 2 marital statuses — married/not married — for full year residents, plus some additional cross-classifications for part-year residents.)

When combined with the information obtained from the census on age and sex characteristics of dependents, the tax record stratified sample data can be compiled into a specially designed "Inference Table", relating the total amount of personal exemptions to the number, sex and age of the dependents. Since the value of personal exemptions changes annually, inference tables are constructed for each tax year. When applied to the records of matched filers, and assuming that dependents move with the tax filer, an estimate of the number, sex and age of migrating dependents is produced.

(3) Adjustment for Coverage

The imputation with respect to dependents results in approximately 90 % of the total population being accounted for by the taxation records. The migration figures require adjustment for the remaining 10 % of population not covered by the taxation system. The adjustment consists of computing the ratio of the population as estimated (or enumerated) to the population of filers and their dependents, and applying the result to the number of migrants. Ratios are calculated for each broad age-sex group, at the sub-provincial census division level.

Potential Errors and Biases

Migration data derived from income tax files are subject to errors and biases other than just those arising from the compilation of the inference algorithm. Three major sources of error are: double counts; noncoverage; and the assumption that dependents move with the tax filers.

(1) Double Counts

It is possible that some persons with low net income may file their own return even though they are claimed by other filers as dependents. A case in point may be wives or husbands working part-time, but whose net income is small enough to make them eligible for inclusion in their spouse's return under partial exemption. Another example is provided by the possible situation of a mother, receiving the child tax credit, who may file a return even though she has no income and is claimed as a dependent. Double counts such as these may explain why, among females in the 18 to 24 and 25 to 44 year age groups, the coverage ratio of filers and dependents combined exceeds 100 % by a considerable margin (Table 4.5).

(2) Noncoverage

Noncoverage arises out of three possible situations: (1) people who neither file an income tax return nor appear as dependents in another filer's return; (2) returns bearing a "non-residential" address;[5] and, (3) on "rejections" due to the inability to match files for two consecutive years (e.g. new filers). The incidence of noncoverage attributable to these factors is shown in Chart 4.3.

(3) Assumption that Dependents Move With Tax Filers

Implicit in the method by which migration data is derived from the income tax files, is the assumption that dependents move with the tax filer. This may be only generally correct, since (particularly in cases of short-term migration, divorce or separation) the filers might not be accompanied by their dependents. Similarly, there may be older child dependents who are still claimed for tax exemption, even though they have established their own households (and migration patterns). In either case, if the filer migrated, the dependents would also be recorded as migrants even though they did not move with the filer. Questions can also be raised about the applicability of the assumption to all groups. Specifically, the rate of migration of particular sub-groups (such as low income adults not captured by

[5] Discounters, i.e., filers who "cash-in" their refund in advance, represent a large part of this category (2/3 in 1984). Under the terms of such an agreement the cheque is redirected to the company and, therefore, the address on the tax return is that of the company. No migration information can be derived in these cases. Another example is professionals — lawyers for instance — who may give their business rather than residential addresses.

the taxation system, or of people counted twice) may indeed be quite different from the rate of migration of the majority of income earners in the same age category. The potential for error arises out of imputing the migration pattern of the specific sub-groups from that of the majority.

Evaluation of Estimates

Family Allowance Data Versus Tax Data

Migration estimates derived from family allowance records and personal income tax records are compared in Table 4.6 (for all migrants) and Table 4.7 (for child migrants). Though the two patterns were found to be highly correlated, they show significant differences in terms of the volume of migration. The estimates of total migration based on taxation data are systematically lower, by a substantial margin, than those based on family allowance (20 % to 30 % lower depending on the year).

One possible explanation for this systematic discrepancy between the two sets of data is that family allowance data are collated monthly, whereas income tax statistics relate to annual periods. All other things being equal, estimates derived from monthly records are expected to exceed those derived from annual records since, for people migrating more than once in a year, the monthly file will record each move (provided that the multiple moves do not occur in the same month), while the annual file will record only one move.

This difference in periodicity could explain part of the differences in the tax and family allowance-based migration levels observed in Tables 4.6 and 4.7. To determine the magnitude of this relationship, an experimental set of child migration data was compiled on an annual basis from a specially designed family allowance programme (F59) for the provinces of Newfoundland, Manitoba, Alberta, British Columbia and the Northwest Territories, for the period 1978-79 to 1981-82.[6] The ratio[7] of annual migration estimates derived from personal income tax files to the migration estimates from family allowance files collated on a 12-month (rather than a monthly) basis, is set out in Table 4.8. When the time period is the same (i.e. one year), the discrepancy between the two sets is attenuated somewhat (6-13 % in the case of children 0-15, and 2-3 % for 0-17).[8] Significant differences still remain, however, especially for children in the 0-15 year age group.

The impact of differences in reference periods (monthly versus annual) may also be examined by comparing the annual net migration estimates. If, within a year, a person moves from province A to province B, and subsequently from

---

[6] Extension of the investigation to other provinces and more recent years was not economically feasible.

[7] Ratios refer to the migration of children aged 0-15 up to 1980-1981 and aged 0-17 in 1981-82.

[8] Family Allowance F59 data for 0-17 in 1981-82 are considered not final and subject to errors.

TABLE 4.6. Ratio Between Revenue Canada Tax Files and Family Allowance M0013 File, for the Number of In-Migrants and Out-Migrants of All Ages: Canada, Provinces and Territories, 1981-82 to 1983-84

Geographic Area	Revenue Canada/M0013		
	In-Migrants		
	1981-82	1982-83	1983-84
Canada	**0.794**	**0.813**	**0.720**
Newfoundland	0.844	0.895	0.769
Prince Edward Island	0.766	0.948	0.852
Nova Scotia	0.837	0.926	0.769
New Brunswick	0.777	0.873	0.740
Quebec	0.720	0.756	0.735
Ontario	0.820	0.873	0.757
Manitoba	0.770	0.864	0.769
Saskatchewan	0.804	0.863	0.712
Alberta	0.792	0.721	0.643
British Columbia	0.777	0.739	0.657
Yukon	0.855	0.760	0.717
Northwest Territories	0.932	0.945	0.958
	Out-Migrants		
Canada	**0.794**	**0.813**	**0.720**
Newfoundland	0.902	0.843	0.782
Prince Edward Island	0.815	0.828	0.735
Nova Scotia	0.856	0.794	0.757
New Brunswick	0.797	0.798	0.690
Quebec	0.914	0.892	0.802
Ontario	0.775	0.757	0.697
Manitoba	0.795	0.763	0.731
Saskatchewan	0.835	0.823	0.750
Alberta	0.743	0.822	0.685
British Columbia	0.718	0.822	0.694
Yukon	0.956	0.836	0.688
Northwest Territories	1.003	1.014	0.917

TABLE 4.7. Ratio Between Revenue Canada Tax Files and Family Allowance M0013 File, for the Number of Child In-Migrants and Out-Migrants: Canada, Provinces and Territories, 1981-82 to 1983-84

Geographic Area	Revenue Canada/M0013		
	In-Migrants		
	1981-82	1982-83	1983-84
Canada	**0.817**	**0.809**	**0.730**
Newfoundland	0.860	0.881	0.767
Prince Edward Island	0.838	0.849	0.800
Nova Scotia	0.863	0.892	0.756
New Brunswick	0.789	0.856	0.715
Quebec	0.732	0.754	0.737
Ontario	0.831	0.872	0.780
Manitoba	0.815	0.845	0.779
Saskatchewan	0.829	0.856	0.719
Alberta	0.830	0.715	0.656
British Columbia	0.775	0.740	0.663
Yukon	0.884	0.773	0.841
Northwest Territories	1.033	0.971	0.961
	Out-Migrants		
Canada	**0.817**	**0.809**	**0.730**
Newfoundland	1.011	0.790	0.783
Prince Edward Island	0.873	0.869	0.732
Nova Scotia	0.877	0.819	0.773
New Brunswick	0.808	0.775	0.692
Quebec	0.903	0.911	0.823
Ontario	0.795	0.756	0.682
Manitoba	0.811	0.780	0.742
Saskatchewan	0.837	0.805	0.781
Alberta	0.766	0.813	0.714
British Columbia	0.758	0.796	0.684
Yukon	0.938	0.924	0.800
Northwest Territories	0.993	1.009	0.945

TABLE 4.8. Ratio Between Revenue Canada Tax Files and Family Allowance F59 File, for the Number of Child In-Migrants and Out-Migrants Aged 0-15 and 0-17: Selected Provinces, 1978-79 to 1981-82

Geographic Area	Migrants Aged 0-15			
	1978-79		1979-80	
	In	Out	In	Out
Newfoundland	0.971	1.035	0.880	0.953
Manitoba	0.880	0.924	0.843	0.878
Alberta	0.907	0.902	0.890	0.885
British Columbia	0.908	0.923	0.877	0.846
Northwest Territories	1.177	1.255	0.922	0.957
Average	**0.912**	**0.936**	**0.880**	**0.881**
	Migrants Aged 0-15		Migrants Aged 0-17	
	1980-81		1981-82	
	In	Out	In	Out
Newfoundland	0.927	0.979	0.983	1.092
Manitoba	0.843	0.836	0.940	0.944
Alberta	0.879	0.849	0.976	0.954
British Columbia	0.934	0.882	N/A	N/A
Northwest Territories	0.950	1.016	1.172	1.188
Average	**0.897**	**0.870**	**0.975**	**0.979**

N/A: Not available.

Note: Until 1981 the taxation records, because of the age structure of the personal exemption in force, only made it possible to generate data on child migration up to age 15, since then the age limit has been extended to age 17. Also data for 1981-82 from family allowance (F-59 annual data set) are unedited data and subject to error.

B to province C, the net result is tantamount to a move from A to C in both annual and monthly data. Yet, the comparison of net child migrants from tax files and family allowance files (M0013), shown in Table 4.9, reveals important discrepancies — as large as 2,062 in the case of Alberta in 1983-84 — between the two sets.

These discrepancies may, in part, result from the following situation: A person moving from province A to province B and then emigrating or dying, will be recorded in a monthly-based record system as a migrant from A to B. In an annual record system, however, the interprovincial migration will not have

TABLE 4.9. Child Net Migration Based on Two Sources

Geographic Area	1981-82		
	FA	Tax	Deviation
Newfoundland	−1,609	−2,091	−482
Prince Edward Island	−102	−131	−29
Nova Scotia	−264	−321	−57
New Brunswick	−670	−642	28
Quebec	−5,188	−5,942	−754
Ontario	−2,217	−829	1,388
Manitoba	−672	−510	162
Saskatchewan	95	16	−79
Alberta	9,518	9,295	−223
British Columbia	1,098	1,139	41
Yukon	56	13	−43
Northwest Territories	−45	3	48
Absolute Sum of Deviations	**N.A.**	**N.A.**	**3,334**
	1982-83		
	FA	Tax	Deviation
Newfoundland	252	507	255
Prince Edward Island	130	93	−37
Nova Scotia	466	791	325
New Brunswick	746	970	224
Quebec	−4,874	−5,484	−610
Ontario	3,942	5,867	1,925
Manitoba	260	651	391
Saskatchewan	879	1,069	190
Alberta	−771	−2,670	−1,899
British Columbia	−486	−1,182	−696
Yukon	−502	−531	−29
Northwest Territories	−42	−81	−39
Absolute Sum of Deviations	**N.A.**	**N.A.**	**6,621**

TABLE 4.9. Child Net Migration Based on Two Sources - Concluded

Geographic Area	1983-84		
	FA	Tax	Deviation
Newfoundland	-797	-667	130
Prince Edward Island	151	181	30
Nova Scotia	1,195	817	-378
New Brunswick	417	397	-20
Quebec	-3,878	-3,830	48
Ontario	10,429	9,945	-484
Manitoba	-33	209	242
Saskatchewan	1,295	543	-752
Alberta	-10,427	-8,365	2,062
British Columbia	2,051	1,076	-975
Yukon	-237	-165	72
Northwest Territories	-166	-141	25
Absolute Sum of Deviations	**N.A.**	**N.A.**	**5,218**

FA: Family Allowance Files.
N.A.:Not applicable.
Source: Statistics Canada, Demography Division

been recorded at all. Such occurrences are probably infrequent, however, and will account for little of the difference between migration levels generated by monthly and annual recording systems. Consequently, the difference in migration levels between the two data sets is due mostly to the imputation of the mobility status and to adjustment for coverage of dependents aged 0-17 years.

Census-Based Error of Closure

Any inference regarding the relative accuracy of the two estimates requires independent criteria of evaluation. One such criterion is the error of closure (the difference between the actual census count and the population estimates at the census date). Error of closure is considered to be a measure of the reliability of an estimating method. The "best" interprovincial migration estimate, therefore, is the one that generates the smallest error of closure, assuming that two successive censuses are consistent in terms of coverage, and that the measures or estimates of the other components of population growth are accurate.

Table 4.10 shows the error of closure for 1981 with respect to both the family-allowance and the taxation-based estimates of migration. The mean absolute error for the 10 provinces is 0.86 % and 0.69 %, respectively, too small a difference to afford any insight into the relative merits of either method. Besides, the error of closure as a criterion of validation rests heavily on the assumption that the population growth components are reliable. Though the provincial figures on births, deaths and immigration are deemed to be reliable, the estimates of emigration are clearly problematic. Furthermore, census counts are subject to undercoverage, the extent of which varies from one census to another. After adjustments of the census counts for undercoverage, the mean absolute error of closure for the provinces becomes 0.90 % and 0.82 %, again too small a difference to be statistically significant.[9] In brief, the evaluation based on the error of closure remains inconclusive. Both types of interprovincial migration estimates yield postcensal estimates that compare satisfactorily with population census counts. However, in terms of error of closure for the 1976-1981 period the estimates based on taxation have a slight edge over those based on family allowance.

Conclusion

Both the family allowance files and personal income tax files have strengths and weaknesses as bases for estimating interprovincial migration. Given the high motivation for filing a change of address as soon as a family moves, the family allowance files yield a fairly reliable estimate of child migration. These files do not, however, provide information about adult migrants, making recourse to other data sources (i.e., census or taxation) necessary. The main advantages of the family allowance data is their timeliness and currency, enabling Statistics Canada to produce reliable preliminary estimates, on a quarterly and annual basis, earlier than would otherwise be possible.

By comparison, the population coverage provided by the taxation files is much more comprehensive and representative. The disadvantage of this source is that the tax data seem to understate child migration. Personal income tax files, however, consistently provide migration data at the provincial, territorial and census division levels, by broad age group and sex, whereas the family allowance data do not. Final annual migration estimates are, therefore, produced from taxation data, though the distribution by quarter must still be based on family allowance data. Evaluation of the accuracy of the data on the age structure of migrants, derived from tax files (which constitutes the basis of the $f_{j,k}$ factors), is not possible with the data currently available.

---

9 When absolute errors of closure are weighted for population size of the provinces, very similar values are obtained.

TABLE 4.10. Error of Closure on June 1, 1981 for Population Estimates Using Different Methods for Estimating Interprovincial Migration

Geographic Area	Population Estimates Using Revenue Canada Tax Migration Data	
	Number	Percent [2]
Newfoundland	9,252	1.63
Prince Edward Island	−64	−0.05
Nova Scotia	2,581	0.30
New Brunswick	914	0.13
Quebec	−19,041	−0.30
Ontario	54,808	0.64
Manitoba	10,961	1.07
Saskatchewan	−3,005	−0.31
Alberta	−53,515	−2.39
British Columbia	804	0.03
Yukon	−458	−1.98
Northwest Territories	−3,237	−7.08
Mean Absolute Error [3]	**N.A.**	**0.69**
	Population Estimates Using Family Allowance Migration Data and $f_{(j,k)}$ Tax Factors [1]	
Newfoundland	13,049	2.30
Prince Edward Island	1,244	1.02
Nova Scotia	3,410	0.40
New Brunswick	3,758	0.54
Quebec	−4,788	−0.07
Ontario	32,162	0.37
Manitoba	8,977	0.87
Saskatchewan	2,711	0.28
Alberta	−59,146	−2.64
British Columbia	−2,050	−0.07
Yukon	478	2.06
Northwest Territories	195	0.43
Mean Absolute Error [3]	**N.A.**	**0.86**

N.A.: Not Applicable.

[1] From 1976 to 1980 Revenue Canada data for children were available for age group (0-15) only. Therefore, the $f_{(j,k)}$ factors were calculated using migrants aged (0-15) and 16+ instead of (0-17) and 18+.

[2] Percent error of closure is calculated by using the following equation:

$$\text{Percent error} = \left(\frac{\text{Estimate} - \text{Census}}{\text{Census}}\right) \times 100$$

The territories have been excluded from the calculation of mean absolute error because they are remote regions and a small difference in absolute numbers represents a large percentage error because of the small population size.

[3] $$\text{Mean absolute error} = 1/10 \left[\sum_{i=1}^{10} \left| \text{percent error in } i^{th} \text{ province} \right| \right]$$

Different strategies for deriving internal migration estimates could be considered for the future. For final estimates, consideration might be given to the development of a *composite* model to produce migration estimates from three different sources of migration data: (a) family allowances as the best source of information for child migration; (b) taxation as probably the best source of information on migration among the working population aged approximately 18-64, and, finally; (c) old age security files, which offer considerable potential for inferring migration among the elderly.

CHAPTER V

EMIGRATION*

There are no direct records on emigrants from Canada. To estimate their numbers, it is necessary to rely on information extracted from Canadian administrative sources, or from records (often incomplete) maintained by recipient countries. For this reason, emigration is the weakest link in the population estimation procedure.

Estimation Methodology

Methods of estimating emigration are hampered by a paucity of reliable data. In this chapter, six different methods for estimating emigration are tested. Of the six, the method based on a combination of family allowance and income tax data stands out as the most promising.

Residual Method

The residual method of estimating emigration at the national level involves the following two steps:

(1) Computation of the change in total population between two successive censuses, and;

(2) Subtraction of the change accounted for by natural increase (births minus deaths) and immigration from the change in total population, the result of which is assumed to be attributable to emigration.

The procedure may be stated as follows:

$$E_{(t-5,\,t)} = [P_t - P_{t-5}] - [B_{(t-5,\,t)} - D_{(t-5,\,t)} + I_{(t-5,\,t)}] \qquad (1)$$

where: E = volume of emigration computed "residually";
P = census population counts;
B = births;
D = deaths;
I = immigrants;
t = last year in which a census was conducted.

Table 5.1 provides a sample calculation for the period 1976-81.

---

* This chapter is based on a background paper prepared by Ronald Raby, Pierre Parent and Anatole Romaniuc.

TABLE 5.1. Residual Emigration in Canada, 1976-1981

(1) 1981 Census Counts	24,341,701
(2) 1976 Census Counts (Adjusted for Vacancy Check [1])	23,052,603
(3) Population Growth (1) - (2)	1,289,098
(4) Births - Deaths + Immigrants	1,566,656
(5) Residual (3) - (4)	−277,558

[1] Unlike the 1981 Census count, the official 1976 Census count was not adjusted for the vacancy check. The unadjusted 1976 Census figure was 22,992,603.

TABLE 5.2. Census Counts Adjusted for Undercoverage, 1976 and 1981

	1976	1981
Census Counts	22,992,603[1]	24,341,701
Coverage Rates as Measured by RRC[2]	.9796318[1]	.9799255
Census Counts Adjusted for Undercoverage	23,470,658	24,840,359
Standard Error	±23,957	±23,833

[1] Not adjusted for vacancy check, as this was included in the undercoverage rate.
[2] The 1976 coverage rate takes account of persons who were missed because their dwellings were misclassified as unoccupied.

TABLE 5.3. Residual Emigration 1976-1981
(After adjustment for census undercoverage)

(1) 1981 Census Counts Adjusted for Undercoverage	24,840,359	±(23,833)[1]
(2) 1976 Census Counts Adjusted for Undercoverage	23,470,658	±(23,957)[1]
(3) Population Growth ($P_{81} - P_{76}$)	1,369,701	±(33,793)[1]
(4) Births - Deaths + Immigration	1,566,656	
(5) Residual (3) - (4)	−196,955	±33,793[2]

[1] Standard error of undercoverage

[2] $\sqrt{(23,957)^2 + (23,833)^2} = 33,793$ = standard error of population growth.

The residual is a "rough-and-ready" measure of emigration. Not only is it affected by the reliability of the available statistics on births, deaths, and immigration, but also by the difference in the undercoverage rates of the two consecutive censuses. [1] An estimate of the accuracy of enumeration in the

[1] John J. Kelly. "Alternative Estimates of the Volume of Emigration from Canada, 1961-71", *Canadian Review of Sociology and Anthropology*, 14(1), 1977, pp.57-67.

1976 and 1981 Censuses was obtained by the Reverse Record Check (RRC), described in Chapter I. The 1976 and 1981 population counts, adjusted for undercoverage, are presented in Table 5.2.

An estimate of the emigrant population, after adjustment for undercoverage, together with the standard error of estimate, is shown in Table 5.3.

Method Based on Foreign Source Data

Data on Canadian emigrants may be obtained through immigration statistics of other countries. The total number of emigrants from Canada can be estimated, using a residual approach, in combination with reliable data from the United States and the United Kingdom.

(1) United States

The quarterly reports produced by the United States Department of Justice, Immigration and Naturalization Service contain information on immigrants to the United States whose "country of last residence" was Canada. This information can be used directly for the number of emigrants leaving Canada for the United States.

(2) United Kingdom

Annual estimates of emigrants from Canada to the United Kingdom are based on the International Passenger Survey (IPS), a stratified sampling of all passengers traversing the sea and air routes between the United Kingdom and other countries. All Canadians surveyed who state that they intend to reside in the United Kingdom for a continuous period of at least 12 months, are considered immigrants to the United Kingdom.

(3) Other Countries

Until recently (see Section 4 below), there have been no records available on Canadians emigrating to countries other than the United States and the United Kingdom. Prior to 1981, estimates of emigrants to other countries were derived residually, as illustrated below, for the period 1971-1976.

$$E^{O}_{t,t+1} = \frac{[P_t - P_{t-5}] - [B_{t-5,t} - D_{t-5,t} + I_{t-5,t}] - [E^{US}_{t-5,t} - E^{UK}_{t-5,t}]}{5} \tag{2}$$

where: $E^{O}_{t,t+1}$ = annual number of emigrants to countries other than USA and UK;

P_{t-5} = the 1971 Census population count, adjusted for undercoverage;

where: P_t = the 1976 Census population count, adjusted for undercoverage;

$B_{t-5,t}$ = number of births, 1971 to 1976;

$D_{t-5,t}$ = number of deaths, 1971 to 1976;

$I_{t-5,t}$ = number of immigrants, June 1971 to May 1976;

$E^{US}_{t-5,t}$ = total number of emigrants from Canada to USA, 1971 to 1976;

$E^{UK}_{t-5,t}$ = total number of emigrants from Canada to UK, 1971 to 1976.

The residual annual number of emigrants to "other countries" for the period 1971-1976 was then used to estimate total emigration during the subsequent period from 1976 to 1981. The total number of emigrants during any period t, t+1 ($E_{t,t+1}$) would be calculated as:

$$E_{t,t+1} = E^{US}_{t,t+1} + E^{UK}_{t,t+1} + E^{O}_{t,t+1} \quad (3)$$

(4) United Nations File on International Immigration

A recent initiative by the Secretariat of Statistics of the Economic Commission for Europe (ECE) gave rise to an exchange of information on migration among 32 European countries, including Canada, the United States and the United Kingdom. As a result, Statistics Canada now receives through the Secretariat, on an annual basis, the number of Canadian immigrants to some of the ECE countries. Efforts are underway to expand this exchange of information on international migration to include the United Nations' network of countries. This information base is still in the developmental stage, but remains a potential data source. Definitional inconsistency is one of the major problems associated with the identification of international migrants. According to the United Nations' definition, long-term immigrants are those nationals and aliens who enter a country with the intent of staying for more than one year. Not all states currently subscribe to this definition.

Method Based on Domestic Sources: Family Allowance and Income Tax Files

Revenue Canada personal income tax files and Health and Welfare Canada family allowance files, described in the previous chapter, both contain address information that can be used to identify emigrants. In the case of the former, the emigrant status is deduced from the returns showing an "out-of-Canada" address one year, and an "in-Canada" address the previous year. In the case of the latter, the emigrant status is derived through the registrees' notification of change of address.[2]

There are important differences between these two data sources. The first difference relates to coverage, with the family allowance domain limited to families with eligible children, while the income tax domain includes all tax-paying individuals and their dependents. The second difference is periodicity. The family allowance files provide figures by month, and are produced in June and December, whereas the tax files contain annual data only, and are produced with a 12 to 15 month lag. To compensate for these differences, and the weaknesses associated with the data, the sets are used jointly in the estimation procedure.

The estimate procedure is similar to the procedure used in Chapter 4 to estimate interprovincial migration:

$$E_j = \left[\frac{E_{j,\,0\text{-}17}}{P_{j,\,0\text{-}17}} \cdot f_j \cdot P_{j,\,18+} \right] + E_{j,\,0\text{-}17} \qquad (4)$$

$$E_c = \sum_{j=1}^{12} [E_j] \qquad (5)$$

where:

E_j = estimated annual number of emigrants from province j;

E_c = estimated annual number of emigrants from Canada;

$E_{j,\,0\text{-}17}$ = the number of emigrants from province j aged 0 to 17 years (inclusive), who are eligible for family allowance (file M0024);

$P_{j,\,0\text{-}17}$ = the number of children in province j who are eligible for family allowance (file M0023);

$P_{j,\,18+}$ = adult population of province j obtained by subtracting the number of children eligible for the family allowance from the total estimated population;

f_j = annual adjustment factor for estimating adult emigration.

---

[2] For estimation purposes, an emigrant is considered to be a Canadian citizen or a landed immigrant who has left Canada on a permanent basis. Excluded are: (1) persons temporarily outside Canada regardless of duration of stay; and, (2) foreign workers returning to their countries who were entitled to family allowance.

The annual adjustment factor (f_j) is calculated as follows:

$$f_j = \frac{\hat{E}_{j,\,18+}}{\hat{P}_{j,\,18+}} \div \frac{\hat{E}_{j,\,0\text{-}17}}{\hat{P}_{j,\,0\text{-}17}} \qquad (6)$$

where:

$\hat{E}_{j,\,18+}$ and $\hat{E}_{j,\,0\text{-}17}$ = estimated numbers of adult and child emigrants from province j, based on Revenue Canada tax files;

$\hat{P}_{j,\,18+}$ and $\hat{P}_{j,\,0\text{-}17}$ = estimated June 1st population of adults and children of province j.

The use of provincial values for f can introduce instability due to the small number of emigrants from some provinces. For this reason, f_j was replaced in the estimating formula by a much more stable, Canada-wide value (f_c). The latter shows a remarkable stability over time (.8698 in 1981-82; .8768 in 1982-83, and; .8846 in 1983-84). The application of f_c to the individual provinces implies invariability in the age structure of emigrants across provinces, however, and could introduce a slight bias in the estimates of the provinces' emigration.

An alternative procedure is to produce a direct estimate of the total number of emigrants from Canada (by formula 4), and then allocate them among the provinces according to a predetermined distribution of emigrants by province. This distribution can be obtained either from family allowance or taxation data. Table 5.4 contains both distributions for the period 1979-1984. Note that in general, the smaller absolute differences between the two distributions, when expressed as a percentage of the family allowance file, are associated with the more populous provinces.

Table 5.5 shows the emigration estimates based on the family allowance proportional distribution for 1981-82, while Table 5.6 provides a direct comparison (index of dissimilarity) between the various methods. The agreement is close, as one would expect, between the three estimation variants based on family allowances. In contrast, the index of dissimilarity is greater between the latter three estimates and the taxation-based estimate.

The Reverse Record Check

Emigrants can be traced through the Reverse Record Check (RRC) conducted after each census to determine the number of persons or households missed by the census. This process is described in Chapter I.

The 1981 Census RRC has traced 468 emigrants for the period 1976-1981 through its sample. When inflated to the total population level, an estimate of 296,725 emigrants for Canada with a standard error of ± 16,021, is produced. The RRC sample, broken down by province, is too small to warrant the use of the RRC data for estimating emigration by province.

TABLE 5.4. Percent Distribution of Emigrants by Province of Origin Based on Two Sources, 1979-1984

Geographic Area	Family Allowance File (1)	Revenue Canada Tax File (2)	Absolute Percentage Difference $\lvert[(1-2)/1] \times 100\rvert$
Canada (Number)	**74,357(a)**	**163,938(b)**	
Canada (Percent)	**100.00**	**100.00**	-
Newfoundland	0.78	1.17	50.0
Prince Edward Island	0.17	0.30	76.5
Nova Scotia	0.91	3.13	244.0
New Brunswick	1.77	2.40	35.6
Quebec	15.82	14.18	10.4
Ontario	47.13	45.13	4.2
Manitoba	3.58	3.93	9.8
Saskatchewan	1.96	2.16	10.2
Alberta	14.30	14.91	4.3
British Columbia	13.32	12.34	7.4
Yukon	0.16	0.13	18.8
Northwest Territories	0.11	0.22	100.0

(a) Number of child emigrants.
(b) Number of emigrants of all ages.

TABLE 5.5. Percent Distribution of Emigrants Based on Alternative Methods Using Family Allowance (FA): Canada, Provinces and Territories, 1981-82

Geographic Area	Method			
	Emigrants Calculated for Canada and Distributed by Province		Emigrants Calculated for Each Province Using Adjustment Factor	
	With FA Data I	With Tax Data II	f_c III	f_j IV
Canada (Number)	**44,367**	**44,367**	**44,823**	**44,660**
Canada (Percent)	**100.00**	**100.00**	**100.00**	**100.00**
Newfoundland	0.75	1.12	0.57	0.66
Prince Edward Island	0.12	0.27	0.11	0.13
Nova Scotia	1.01	3.33	0.95	0.96
New Brunswick	2.17	3.08	1.93	1.95
Quebec	15.15	13.60	15.38	15.67
Ontario	45.98	42.93	46.68	45.56
Manitoba	3.60	4.02	3.47	3.65
Saskatchewan	1.89	2.01	1.73	1.82
Alberta	15.61	16.59	14.94	15.15
British Columbia	13.50	12.67	14.06	14.24
Yukon	0.15	0.16	0.13	0.15
Northwest Territories	0.07	0.22	0.05	0.06

TABLE 5.6. Index of Dissimilarity [1] in Percent

Method	Method		
	II	III	IV
I	5.43	1.48	1.32
II	-	6.91	6.27
III	-	-	1.11

[1] The index of dissimilarity has a range of 0 to 100, and is defined as one-half the sum of the absolute differences between two percent distributions.

TABLE 5.7. Estimated Number of Emigrants from Canada Based on Different Methods, 1976-1981

Methods	Estimate	Standard Error
Residual (Unadjusted for Census Undercoverage)	277,558	N.A.
Residual (Adjusted for Census Undercoverage)	196,955	± 33,793
Foreign Sources (USA & UK) & Assumption of 48,000 Migrants to All Other Countries	371,655	N.A.
Tax Files (Obtained Directly, Independently of Family Allowance Data)	207,420	N.A.
Family Allowance[1]	278,624	N.A.
Reverse Record Check	296,724	± 16,021

N.A.: means not available.
[1] Estimated using f_c for all provinces.

Evaluation of Estimates

Emigration estimates for the period June 1, 1976 to June 1, 1981 produced by means of the six procedures described in the foregoing, are compared in Table 5.7.

Evaluating the quality of emigration estimates is a difficult enterprise. The most plausible "objective criteria" for evaluation appear to be: (1) the consistency between alternative estimates, and; (2) the "error of closure", or the difference between census counts and estimates of population at the same reference date. In the first case, the underlying assumption is that if independently derived estimates tend to agree, then they provide a reliable approximation of emigration. Underlying the error of closure approach is the rationale that the "best" emigration estimate is the one which tends to minimize the difference between the population estimates and the census population counts.

Of the six estimates presented in Table 5.7, the family allowance and residually (unadjusted for undercoverage) derived estimates best satisfy the consistency criterion, with estimates of 278,624 and 277,558 emigrants, respectively. They also satisfy what may be called the "central tendency" criterion in that they most closely approximate the arithmetic average (271,493) of the six estimates. The extreme values, 371,655 from foreign sources and 207,420 from the taxation file as well as the 198,955 derived residually after adjustment for census undercoverage, are all implausible. The taxation file seems to understate the actual number of emigrants, particularly among children. As shown in Table 5.8, the taxation file yields significantly lower estimates of emigration of children (roughly 30 % to 35 %), compared to that based on family allowance data.

Of the six types of emigration estimates discussed, four can be evaluated against the error of closure. At the national level, the family allowance estimate yields the best fit for the period 1976-1981, as shown in Table 5.9. The Reverse Record Check (RRC) has the next smallest error of closure. The latter is rejected, however, on two grounds: (a) it can only be carried out every five years after the census, and; (b) it is unreliable at provincial levels because of the small RRC sample size. The remaining two estimates — taxation and foreign files — have the largest absolute error of closure of the four estimates.

In conclusion, the method based on the combined use of family allowance data for child emigration, and taxation data for the child-adult migration ratio, stands out as the best estimation procedure, and has been chosen for estimating annual emigration flows from 1981 onward.

TABLE 5.8. Comparison of Estimates of the Number of Emigrant Children Based on Income Tax Files: Canada, 1980-81 to 1982-83

Year	Income Tax (1)	Family Allowance (2)	Ratio (1) ÷ (2)
1980-81	8,910	13,067	0.68
1981-82	9,493	13,679	0.69
1982-83	9,601	14,811	0.65

TABLE 5.9. Error of Closure[1] on June 1, 1981, Resulting from Emigration Estimates Based on Different Techniques, 1976-1981

Type of Emigration Estimates	Error of Closure	
	Number	Percent
Foreign Administrative Files	+94,097	0.39
Tax Files	−70,138	−0.29
Reverse Record Check 1981	+19,166	0.08
Family Allowance[2]	+1,006	Negligible

[1] Postcensal estimates minus census counts (i.e. estimates using new methods).
[2] Estimated as stated in Table 5.6.

CHAPTER VI

LOCAL AREA POPULATION*

There are two types of postcensal population estimates produced for local areas (Census Divisions and Census Metropolitan Areas): (i) preliminary estimates based on the regression method, and; (ii) final estimates based on the component method. The former are produced within six months of the reference date, whereas the latter do not become available until approximately 20 months after the fact.

Estimation Methodology

The component and regression estimates are produced independently and, thus, could conceivably lead to inconsistent results for a given reference date. Hence a "regression-nested method" has been devised to tie these two sets of estimates together. Regression estimates are used to measure the change from one point in time to another, and this change is added to the previous year's component estimate to derive the regression-nested estimates (see Table 6.1).

The component method used for local area estimates is the same as that used to estimate the total population of Canada, the provinces and territories, discussed in Chapter I. This chapter focuses on the regression method developed to produce the preliminary postcensal population estimates for census divisions and census metropolitan areas.

TABLE 6.1. Methodology for Preliminary Population Estimates (Regression-nested) for Census Divisions

Time	Regression Estimate	Component Estimate[1]	Preliminary Estimate
t + 1	$\hat{P}_{t+1}$	$\hat{P}'_{t+1}$	$\hat{P}_{t+1}$
t + 2	$\hat{P}_{t+2}$	$\hat{P}'_{t+2}$	$\hat{P}'_{t+1} + [\hat{P}_{t+2} - \hat{P}_{t+1}]$
t + 3	$\hat{P}_{t+3}$	$\hat{P}'_{t+3}$	$\hat{P}'_{t+2} + [\hat{P}_{t+3} - \hat{P}_{t+2}]$
t + 4	$\hat{P}_{t+4}$	$\hat{P}'_{t+4}$	$\hat{P}'_{t+3} + [\hat{P}_{t+4} - \hat{P}_{t+3}]$
t + 5	$\hat{P}_{t+5}$	$\hat{P}'_{t+5}$	$\hat{P}'_{t+4} + [\hat{P}_{t+5} - \hat{P}_{t+4}]$

[1] The method uses census counts as a base population plus births and deaths from Vital Statistics Records and migration data derived from Revenue Canada Taxation Files.

---

* This chapter is based substantially on the work of Ravi B.P. Verma, K.G. Basavarajappa and Rosemary Bender, in "New Approaches to Methods of Estimating the Population of Census Metropolitan Areas and Census Divisions", unpublished working paper, Statistics Canada, Demography Division, Ottawa, 1982.

Regression Method

The regression method, employing either the ratio-correlation or the difference-correlation technique, is used to estimate any event (dependent variable) using a set of symptomatic indicators as predictors (independent variables). The equation expressing this relationship for the j^{th} small area is given by:

$$\Delta y_j = \alpha + \beta_1 \Delta x_{1j} + \beta_2 \Delta x_{2j} + \ldots + \beta_n \Delta x_{nj} + \varepsilon \qquad (1)$$

where:

Δy = the vector of change in proportional values of the dependent variable between census years t and t − 5;

Δx = the vector of change in proportional value for the specific independent variable (symptomatic indicator);

α = the regression intercept;

β = the regression coefficient for the specific independent variable;

ε = the vector of stochastic errors, such that $E(\varepsilon) = 0$ and $E(\varepsilon\varepsilon') = \sigma^2$.

In the ratio-correlation regression method,

$$\Delta y_j = \frac{P_{j,t}}{P_t} \div \frac{P_{j,t-5}}{P_{t-5}} \qquad (2)$$

and in the difference-correlation method,

$$\Delta y_j = \frac{P_{j,t}}{P_t} - \frac{P_{j,t-5}}{P_{t-5}} \qquad (3)$$

where: $y_j = P_{j,t} / P_t$

and;

t and t − 5 = the two census years from two consecutive censuses;

P = the population of a province (or larger region);

P_j = the population of the j^{th} small area, such that $\Sigma P_j = P$;

and the derivation of the Δx's can be inferred from that given for the Δy's.

Equation 1 is fitted by the Least Squares method. Problems of multicollinearity and homogeneity of variance of the error term (ε) are taken into account when necessary, before the estimates of the regression coefficients are derived. A weighted regression method was adopted in order to control for heteroscedasticity; in this procedure, the data set is transformed with calculated weights, such that a random error term (ε), with constant variance, is obtained.

The Goldfield-Quandt procedure has been used to test the assumption of homoscedasticity,[1] whereas ridge regression is used to control for multicollinearity. In this procedure, estimates of β coefficients are obtained by adding a small value (K = .04) to the diagonal of the correlation matrix (X'X).

Knowing the values of α, the β's, of P_{t+i}, and the value of the Δ x's at time t + i, the population of a small area P_j at time t + i can be calculated, for any i^th year, where i ranges from 1 to 5.

In ratio-correlation, the population estimates are given by:

$$\hat{P}_{j,\,t+i} = [\, y_{j,\,t} \cdot \hat{\Delta} y_j \,] \cdot \hat{P}_{t+i} \qquad (4)$$

In difference-correlation, the population estimates are given by:

$$\hat{P}_{j,\,t+i} = [\, y_{j,\,t} + \hat{\Delta} y_j \,] \cdot \hat{P}_{t+i} \qquad (5)$$

where:

$\hat{P}_{j,\,t+i}$ = estimated population for small area j at time t + i;

$\hat{P}_{t+i}$ = independent estimate of population for the larger area containing all the small areas at time t + i;

$y_{j,\,t}$ = proportion of the population of small area j with respect to the larger area, according to the census counts at time t;

$\hat{\Delta} y_j$ = estimate of Δy_j, the change in proportion of the population of small area j with respect to the larger area, between time t and t - h as derived by the regression model.

Neither the ratio-correlation nor the difference-correlation method uniformly or routinely outperforms the other.[2] A choice between the two is thus con-

[1] J. Johnston, *Econometric Methods*, McGraw-Hill, New York, 1963, p. 219.

[2] M. Mandell, and J. Tayman, "Measuring Temporal Stability in Regression Model of Population Estimation", *Demography,* Vol. 19, No. 1, February 1982, pp. 135-146.

O'Hare, W., "Report on a Multiple Regression Method for Making Population Estimates", *Demography,* Vol. 13, No. 3, August 1976, pp. 369-380.

Ravi B.P. Verma, K.G. Basavarajappa and Rosemary K. Bender, "The Regression Estimates of Population for Sub-Provincial Areas in Canada", *Survey Methodology,* Vol. 9, No. 2, December 1983, pp. 219-240.

Ravi B.P. Verma, K.G. Basavarajappa and Rosemary K. Bender, "Estimation of Local Area Population: An International Comparison", *Proceedings of the Social Statistical Section, American Statistical Association,* Washington, D.C., 1984, pp. 324-329.

tingent upon a thorough evaluation of the results to date. The estimation procedure depends, in certain respects, on whether it applies to census divisions or census metropolitan areas, as discussed below.

Census Divisions (CDs)

The specifications for the regression models for each province are based on census division data. These are presented in Table 6.2, and the regression coefficients in Table 6.3. The regression models for each province utilize the best available symptomatic indicators of population change, namely; the number of family allowance recipients aged 1-14 for most provinces; the population registered in provincial health insurance programs (Saskatchewan and Alberta), and; the number of hydro accounts (British Columbia). The main reason for using different regression models is to maximize the accuracy of population estimates by taking advantage of administrative files containing data on specific local areas. The regression method (as well as the variables) selected is the one which results in the lowest mean absolute error, as defined in Table 6.2. The sum of the regression estimates for census divisions within a specific province is adjusted to the corresponding provincial total obtained by the component method.

Census Metropolitan Areas (CMAs)

The CMA population estimates are based on aggregating the census division regression-nested estimates[3] for all CMA convergent CDs (i.e. all CDs whether located entirely within CMAs, or whose boundaries intersect CMA boundaries). Based on the CD and CMA populations (enumerated, or estimated) for the previous year, the population of a CMA is expressed as a ratio of the sum of the populations of the "convergent" CDs. This ratio (assumed to remain constant for two successive years), is then applied to the sum of the reference date CD population estimates, providing an estimate of population for the CMA at the same reference date. This operation is summarized in the following equation:

$$\hat{P}^{cm}_{t+1} = \frac{P^{cm}_t}{\sum_{cd} P^{cd}_t} \cdot \sum_{cd} P^{cd}_{t+1} \tag{6}$$

where:

$\hat{P}^{cm}_{t+1}$ = estimated population of CMA in year $t + 1$;

P^{cm}_t = estimated or enumerated CMA population in year t;

---

[3] Bender R. and R. Verma, "Translation for Converting Demographic Data Between Overlapping Subprovincial Areas in Canada", *Proceedings of the Social Statistics Section of the American Statistical Association,* Washington, D.C. 1983, pp. 518-521.

TABLE 6.2. Specifications of the Regression Method for Estimating Post-1981 Total Population for Census Divisions (CDs) Within Provinces, and for Census Metropolitan Areas (CMAs)

Geographic Area	Number of CDs or CMAs [1]	Type [2]	Model Period	Symptomatic Indicator(s)	Test 1981 MAE
Newfoundland and Prince Edward Island	13	RC	1976-81	F	1.27
Nova Scotia	18	RC	1971-76, 1976-81	F	1.50
New Brunswick	15	RC	1976-81	F	1.30
Quebec	76	RC	1976-81	F	1.81
Ontario	53	RC	1976-81	F	1.99
Manitoba	23	WDC	1971-76, 1976-81	F	3.13
Saskatchewan	18	DC	1976-81	RP	0.62
Alberta	15	WRC	1976-81	F, RP	1.89
British Columbia	29	WDC	1971-76, 1976-81	F, Hydro	2.14
CMAs	24	RC AGG	1976-81	F	2.30 1.30

Note: F = family allowance recipients aged 1-14 years.
RP = reference population obtained from health insurance files.
Hydro = number of hydro connections.

$$\text{MAE} = \text{mean absolute error} = \frac{1}{n}\sum \left| \frac{\hat{P}_j - P_j}{P_j} \right| \cdot 100$$

$\hat{P}_j$ = estimated population of census division j.
P_j = census population of census division j.
n = number of census divisions in a given province.
RC = ratio-correlation.
WDC = weighted difference-correlation.
WRC = weighted ratio-correlation.
DC = difference-correlation.
AGG = CMA estimates obtained by aggregating the appropriate CD estimates, with the exception of Calgary, where the estimate is based on an annual census count conducted by the city.

[1] Excluding Yukon and Northwest Territories.

[2] For a description of the types of regression methods, the readers are referred to the paper by W. O'Hare (see footnote 2).

TABLE 6.3. Regression Coefficients[1], for Estimating Post-1981 Population of Census Metropolitan Areas (CMAs), and Provinces (from Census Division Level Data)

Geographic Area	Model Type	Period	α	β_1	β_2
Newfoundland and Prince Edward Island	RC	1976-81	.478	.514	N.A.
Nova Scotia	RC	1971-76, 1976-81	.467	.526	N.A.
New Brunswick	RC	1976-81	.503	.498	N.A.
Quebec	RC	1976-81	.385	.621	N.A.
Ontario	RC	1976-81	.256	.741	N.A.
Manitoba	WDC	1971-76, 1976-81	.000	.609	N.A.
Saskatchewan	DC	1976-81	.000	1.086	N.A.
Alberta	WRC	1976-81	.088	.460	.476
British Columbia	WDC	1971-76, 1976-81	.000	.376	.606
CMA	RC	1976-81	.139	.862	-

N.A.: Not applicable since only one symptomatic indicator is used.
[1] β_1 = regression coefficient of first symptomatic indicator.
β_2 = regression coefficient of second symptomatic indicator.

P^{cd}_{t+1} = estimated population in year t + 1 for a census division, either wholly contained within the CMA, or whose boundary intersects the CMA boundary (estimate, obtained by the regression-nested procedure);

P^{cd}_{t} = estimated or enumerated population in year t for a census division either wholly contained within the CMA, or whose boundary intersects the CMA boundary.

At the census metropolitan area level, this method has been found to be superior to the regression method (see Table 6.2).

Evaluation of Methods

A "multiple-model" framework seems to be the most appropriate for evaluating different estimation methods. The methods tested have included:

(1) Two types of component method using migration estimates derived from school enrolment and tax file data;

(2) A vital rates method;

(3) A ratio method using data from provincial administrative files;

(4) A proportional allocation method based on family allowance recipients, and;

(5) Six types of regression method (ratio-correlation, weighted ratio-correlation, ridge weighted ratio-correlation, difference-correlation, weighted difference-correlation, and ridge weighted difference-correlation).

The accuracy of the procedures and the usefulness of the symptomatic indicators is determined by the mean absolute error (MAE), defined as the mean of the percentage absolute difference between the estimated and the census populations. The procedures and the symptomatic indicators chosen for producing estimates during the post-1981 period were those which gave timely estimates with minimum MAE. The MAE (based on census division data) in all provinces, except Manitoba and British Columbia, is less than 2 % (Table 6.2).

Table 6.4 compares the estimates obtained by the regression and the component methods, with the 1981 Census counts. For Canada as a whole, the method that seems to produce the most accurate results is the regression-nested method (mean absolute error 1.7 %). Both the multiple and simple regression methods are observed to be less accurate than either the component or the regression-nested procedure. This is also true for all the provinces except Saskatchewan where, based on data from health insurance files as the symptomatic indicator, the MAE is lower for the regression method than for either the component or the regression-nested method. In 5 out of 10 provinces, the regression-nested method is slightly more accurate than the component method.

The relative accuracy of each of the alternative methods was also assessed by means of the paired t-test. For Canada, as well as for the provinces of Ontario and Quebec, the difference between the estimates obtained from the standard regression (i.e. not nested) and the component methods is statistically significant at the 1 % level. By contrast, the difference between the regression-nested and the component methods is not statistically significant. Similar t-test results are obtained when the means of absolute error are weighted in accordance with population size.

TABLE 6.4. Percent Mean Absolute Error of Population Estimates for Census Divisions, June 1, 1981

Geographic Area	Number of Census Divisions [1]	Regression Method [2]	Regression-Nested	Component Method
Total	**260**	**2.55**	**1.72**	**1.80**
Newfoundland and Prince Edward Island	13	1.36	0.67	1.00
Nova Scotia	18	1.64	1.27	1.07
New Brunswick	15	1.59	1.05	1.06
Quebec	76	3.10	1.63	2.02
Ontario	53	2.17	1.26	1.21
Manitoba	23	3.33	2.57	2.58
Saskatchewan	18	1.43	1.96	2.10
Alberta	15	4.45	2.84	2.39
British Columbia	29	2.45	2.50	2.39

[1] Excluding Yukon and Northwest Territories.
[2] The method uses as symptomatic variables, health insurance data for Saskatchewan and family allowance recipients for the other provinces. The model period for all provinces is 1971-1976, using weighted ratio correlation for Alberta, weighted difference correlation for British Columbia, and ratio correlation for all other provinces.

Temporal Stability

The temporal stability of the regression, component and regression-nested methods was assessed by means of an index of dissimilarity, computed for the years 1977 through 1981 (Table 6.5). This index, defined as one-half the sum of the absolute differences between two percentage distributions of the populations of census divisions in each province, has a range of 0 to 100. The gap between the regression-direct and component distributions (A) increases over time for most provinces, as does the disparity between the regression-direct and regression-nested distributions (C). The two methods producing the least disparity are the regression-nested and the component methods (B). It should be noted, however, that the component and the standard regression methods are independent. Accordingly, the disparity (A) between these two methods may be expected to vary more widely than between the regression-nested and the component methods, which are not independent.

The difference between the regression-direct and the component estimates tends to be larger than the other two differences. The assumption that the vector of regression coefficients is invariant from the immediately preceding period ($t - 5$, t) to the subsequent period (t, $t + 5$), is often questionable. In practice, this invariance may not hold true over time, both because of struc-

TABLE 6.5. Percent Index of Dissimilarity [1] Between Three Sets of Postcensal Estimates for Census Divisions [2] (Regression, Regression-nested, Component), 1977-1981

Geographic Area		1977	1978	1979	1980	1981
Newfoundland	A.	0.17	0.33	0.41	0.34	0.51
	B.	0.17	0.19	0.19	0.13	0.13
	C.	N.A.	0.17	0.34	0.35	0.41
Prince Edward Island	A.	0.17	0.26	0.25	0.51	0.51
	B.	0.17	0.08	0.19	0.02	0.24
	C.	N.A.	0.17	0.26	0.52	0.26
Nova Scotia	A.	0.29	0.53	0.60	0.63	0.64
	B.	0.29	0.30	0.18	0.23	0.19
	C.	N.A.	0.27	0.52	0.59	0.63
New Brunswick	A.	0.52	0.38	0.46	0.71	0.48
	B.	0.52	0.48	0.44	0.52	0.37
	C.	N.A.	0.53	0.38	0.45	0.70
Quebec	A.	1.02	0.64	0.81	0.99	1.13
	B.	1.02	0.72	0.27	0.57	0.54
	C.	N.A.	1.05	0.66	0.80	0.98
Ontario	A.	1.69	0.58	0.70	0.99	0.94
	B.	1.69	1.75	0.31	0.49	0.56
	C.	N.A.	1.67	0.55	0.71	0.96
Manitoba	A.	0.21	0.39	0.60	0.70	0.80
	B.	0.21	0.26	0.26	0.21	0.19
	C.	N.A.	0.20	0.42	0.59	0.70
Saskatchewan	A.	0.37	0.52	0.53	0.70	0.78
	B.	0.37	0.18	0.26	0.25	0.18
	C.	N.A.	0.38	0.51	0.55	0.68
Alberta	A.	0.45	0.45	0.57	0.89	1.18
	B.	0.45	0.21	0.27	0.41	0.36
	C.	N.A.	0.44	0.43	0.56	0.86
British Columbia	A.	0.39	0.45	0.76	0.95	0.93
	B.	0.39	0.32	0.41	0.23	0.29
	C.	N.A.	0.37	0.43	0.76	0.94

N.A.: Not applicable since population estimates are not produced for census years.

[1] Index of dissimilarity (ID) between estimates P_j' and P_j for a province with n census divisions and total provincial population P is calculated as:

$$ID = \frac{1}{2} \sum_{j=1}^{n} \left[\frac{P_j - P_j'}{P} \right] \cdot 100$$

[2] Index of dissimilarity:

A. regression vs. component estimates.
B. regression-nested vs. component estimates.
C. regression vs. regression-nested estimates.

TABLE 6.6. Comparison of the Accuracy of the Regression Methods for the Model Periods, 1971-1976 and 1976-1981

Geographic Area	Regression		Model 1971-1976		Model 1976-1981
	Type	Indicator(s)[1]	Test 1976[2] MAE	Test 1981[3] MAE	Test 1981[4] MAE
Census Division Total			**1.96**	**2.54**	**2.04**
Newfoundland and Prince Edward Island	RC	F	1.6	1.4	1.3
Nova Scotia	RC	F	1.8	2.0	1.6
New Brunswick	RC	V,F	2.0	1.0	0.9
Quebec	RC	V,F	1.4	2.3	1.8
Ontario	RC	V,F	2.0	2.5	2.1
Manitoba	RC	F	1.9	3.3	3.5
Saskatchewan	RC	RP	1.5	1.3	0.7
Alberta	RC	F	3.1	4.6	4.2
British Columbia	WDC	F	3.1	4.0	2.3

[1] F = family allowance recipients aged 1-14 years.
V = vital events (births + deaths).
RP = reference population obtained from health insurance files.
RC = ratio-correlation.
WDC = weighted difference-correlation.
[2] 1971-76 data used to estimate 1976.
[3] 1971-76 data used to estimate 1981.
[4] 1976-81 data used to estimate 1981.
Source: Statistics Canada, Demography Division.

tural changes in the underlying relationships of the variables, as explained below, and also because of the improvement in the quality of the symptomatic indicators over time. Consequently, the model may work well for one time period, but predict poorly for the subsequent period.

The Effects of Structural Changes

In order to examine the effects of structural changes on the accuracy of the regression estimates, the mean absolute errors of the 1981 estimates derived from a model based on the 1971-1976 period are compared with those from a model based on the 1976-1981 period (Table 6.6). In general, the errors are smaller when the equations are based on the more recent time period, particularly for Saskatchewan and British Columbia, where mean absolute error is reduced by nearly 50 % (from 1.3 % to 0.7 %) and 40 % (from 4.0 % to 2.3 %), respectively. In practice, however, data for 1976-1981 would not be available to produce the estimates for 1977 to 1981, and thus the errors based on the model period 1971-1976 are more realistic.

As one of its main premises, the regression model assumes that the relationship between the variations in symptomatic indicators (vital events and family allowance recipients) and population size (used to derive the regression equations), will remain constant for the subsequent period of estimation. However, this is not always the case. For example, the characteristics of migrants to the western provinces between 1976 and 1981 may differ considerably from the characteristics of migrants between 1971 and 1976. As family allowances are limited to families with children of eligible age, (0-17 years) movements of single persons, and families without children, are not captured by the family allowance indicator. The family allowance indicator derived from 1971-1976 is not properly representative of the changes in migration that occurred in 1976-1981. Thus, in this way, structural changes contribute to the mean error differences observed in 1981, computed from models of each time period, 1971-1976 and 1976-1981.

Part of the difference in the mean absolute errors between 1976 and 1981 is due to problems in the creation of the family allowance recipient data set. The number of family allowance recipients at the census division level is calculated by converting postal codes to standard geographic codes. In 1976, there were problems in the conversion file because of missing and overlapping postal codes. The percentage of family allowance files with missing codes was particularly high in the Atlantic provinces and in Ontario. By 1981, the magnitude of the problem of missing postal codes had been reduced in all provinces. Thus, the improvement in the quality of the family allowance indicator between 1976 and 1981 may have also had an effect on the quality of the regression coefficients for the 1976-1981 period.

Table 6.7 presents the mean absolute errors with respect to the 1981 population estimates for both "types" of CMA; CMAs as statistically autonomous

TABLE 6.7. Percent Mean Absolute Error of Closure (MAE) and Percent Index of Dissimilarity (ID) for Three Methods of Generating Population Estimates: 1981 CMAs and Aggregated CDs

Method	CMA as autonomous unit				Aggregated CDs			
		ID Versus Method				**ID** Versus Method		
	MAE	I	II	III	**MAE**	I	II	III
I Regression[1]		–	1.15	1.09		–	0.86	0.54
II Component	1.47	1.15	–	0.98	1.20	0.86	–	0.48
III Regression-nested	2.21	1.09	0.98	–	1.03	0.54	0.48	–

[1] Regression F based on family allowance recipients aged 1-14 years for 1971-1976 with ratio-correlation.

Source: Statistics Canada, Demography Division, May, 1985.

units, and CMAs as aggregates of census division estimates. Estimates were calculated using the standard regression, regression-nested and component methods. Population estimates obtained by aggregating census division estimates show considerably lower errors than those obtained for CMAs as autonomous units. Among the three methods, it is the component method which produces the lowest error for estimation of CMA population without census division aggregation. With aggregation, the regression-nested procedure produces superior results.

A paired comparison of the three methods, based on the indices of dissimilarity in Table 6.7, shows that the differences among the estimates are small and that both the regression-nested and the component methods provide very similar results.

The regression method, using family allowance recipients and/or other variables as symptomatic indicators, provides estimates which are timely (i.e. with a delay of only about 4 months), though slightly less accurate. On the other hand, the regression-nested method generates estimates which are both timely and accurate, the mean absolute error being roughly comparable to that of the component method.

Evaluation of Estimates

Both the regression and component based estimates of local area populations for the postcensal years 1982-1986 will be subject to a margin of error which may vary from one area to another. Tables 6.8 and 6.9 give the distribution of census divisions in 1981 ranked by the magnitude of the errors of closure for 1981, expressed as a percentage of the Census counts. It should be emphasized that the comparisons based on these errors of closure must be interpreted with caution. The magnitude of the errors observed for the period 1976-1981 could vary for the period 1981-1986. It should also be noted that the errors in the 1981 estimates have cumulated over a five-year period.

The two main sources of error are census undercoverage and errors in the estimation of components of population change. Boundary changes may also be a factor.[4] Special studies have established that the undercoverage rates for Canada in the 1976 and 1981 Censuses averaged about 2 %, but were not equally distributed throughout the country.[5] The undercoverage rates certainly vary among census metropolitan areas (Table 6.10). (Although the studies did not examine census divisions, there are strong indications that the rates also

[4] The CMA and CD boundaries used to calculate the errors are those in effect for the 1981 Census, as the differences between the 1976 and 1981 boundaries were minor. It should, however, be noted that the boundaries for the period 1981-1986 are those of the 1981 Census. Therefore, any boundary changes incurred in the administrative divisions by the various provinces subsequent to the 1981 Census are not taken into account.

[5] See notes to Table 6.10.

TABLE 6.8. Distribution of Census Divisions (CDs) by Province, Magnitude of Error of Closure and Overall Mean Absolute Error of Closure (MAE) for the 1981 Regression-Nested Estimates[1]

Geographic Area	Absolute Error of Closure[2]						
	Number of Census Divisions	Less than 1.0%	1.0%-1.9%	2.0%-2.9%	3.0%-3.9%	4.0% and over	MAE (in Percent)
Total CDs[3]	**260**	**97**	**81**	**44**	**17**	**21**	**1.72**
Newfoundland	10	6	4	0	0	0	0.72
Prince Edward Island	3	3	0	0	0	0	0.50
Nova Scotia	18	9	5	3	0	1	1.27
New Brunswick	15	9	4	1	0	1	1.05
Quebec	76	22	27	13	8	6	1.63
Ontario	53	28	13	9	2	1	1.27
Manitoba	23	8	4	5	1	5	2.56
Saskatchewan	18	2	6	7	2	1	1.96
Alberta	15	0	7	2	3	3	2.84
British Columbia	29	10	11	4	1	3	2.49

[1] Estimates calculated using the 1971-1976 model.
[2] The absolute error of closure is expressed as a percent of the population enumerated.
[3] Excluding Yukon and Northwesi Territories.

TABLE 6.9. Distribution of Census Divisions (CDs) by Province and Census Metropolitan Areas (CMAs), Magnitude of Error of Closure, and Overall Mean Absolute Error of Closure (MAE) for the 1981 Component Estimates

Geographic Area	Absolute Error of Closure[1]						
	Number of CDs or CMAs	Less than 1.0%	1.0%-1.9%	2.0%-2.9%	3.0%-3.9%	4.0% and over	MAE (in Percent)
Total CDs	**265**	**115**	**77**	**36**	**12**	**25**	**1.77**
Newfoundland	10	4	1	4	1	0	1.59
Prince Edward Island	3	1	2	0	0	0	0.96
Nova Scotia	18	11	6	0	0	1	1.06
New Brunswick	15	12	2	0	0	1	1.32
Quebec	76	35	22	10	4	5	1.60
Ontario	53	30	14	6	1	2	1.28
Manitoba	23	6	4	6	1	6	1.28
Saskatchewan	18	2	7	5	2	2	2.08
Alberta	15	4	6	1	1	2	1.83
British Columbia	29	10	12	2	1	4	2.32
Yukon	1	0	0	1	0	0	2.12
Northwest Territories	4	0	1	1	0	2	7.11
Total CMAs	**24**	**9**	**12**	**2**	**0**	**1**	**1.21**

[1] The absolute error of closure is expressed as a percent of the population enumerated.

vary from one census division to another.) The error of estimation, by comparison, is due in part to: the weakness inherent in assuming the invariance of the regression coefficients in estimating population change; shortcomings in the quality of the input data, and; to the uniqueness of symptomatic indicators for each province. Unfortunately, save for comparing these estimates to estimates developed from alternative data sources, there is no direct way to measure the error of estimation.

TABLE 6.10. Estimated Population Undercoverage Rates in Percent for Census Metropolitan Areas (CMAs), 1976 and 1981

Geographic Area	1976		1981	
	Estimated Rate	Standard Error	Estimated Rate	Standard Error
All CMAs	**2.31**	**0.14**	**2.16**	**0.14**
CMA Parts				
Urban Core	2.77	0.22	2.22	0.15
Urban Fringe	1.90	0.21	1.99	0.83
Rural Fringe	1.76	0.46	1.05	0.40
Specific CMAs				
Montreal	3.74	0.40	2.09	0.30
Toronto	2.32	0.34	2.77	0.26
Vancouver	2.86	0.45	2.52	0.38
Other	1.51	0.18	1.83	0.21

Sources: *Quality of Data*, Series I: *Sources of Error Coverage*, 1976 Census of Canada, Catalogue No. 99-840, Statistics Canada, March 1980.
Data Quality - Total Population, 1981 Census of Canada, Catalogue No. 99-904, Statistics Canada, (forthcoming).

CHAPTER VII

FAMILY ESTIMATES*

The publication of census family estimates dates back to 1948, when Statistics Canada released its first set of family estimates.[1] Several methodological revisions have been made in the interim, the latest in 1977,[2] with the additional component of common-law families having been introduced in 1982. The methodology consists of two steps:

(a) The determination of the total number of families by means of the component method, involving the addition of newly formed families to, and subtraction of dissolved families from, the number of families at base year;

(b) The breakdown of the total number of families by sex and age of parents, and further by family size and children's age, accomplished on the basis of data from the two previous censuses.

For the purpose of generating estimates, the definition of family is the same as that used for the census. As such, a census family can be any of the following:

(1) A husband and a wife living in the same dwelling, with or without children (regardless of their age but who have never married), or;

(2) A lone-parent of any marital status, living in the same dwelling with one or more children (regardless of their age but who have never married), or;

(3) Persons living in common-law relationships, regardless of their legal marital status, but who are considered to form a husband-wife family (as in 1, above).

Estimation Methodology

Total Number of Families

The change in the number of families between any two points in time (t, t + 1) is determined by events which result in family formation or dissolution. These events, classified by their impact on the number of families, are

---

* This chapter is based on an earlier working paper by Yolande Lavoie on Family Estimates, subsequently updated by Y. Edward Shin.

1 *Estimates of Households and Families in Canada,* 1947-1949, Reference No. 7-911-0, Dominion Bureau of Statistics.

2 *Estimates of Families for Canada and the Provinces,* 1977 and 1978, Catalogue No. 91-204, Statistics Canada.

TABLE 7.1. Events and Hypotheses Used to Reconstruct Changes in the Number of Census Families

Events	Hypothesized Effect of Events on Number of Census Families		
	Increase	No change	Decrease
Marriages and common-law unions [1]	If neither spouse is a family head	If one spouse is a family head	If both spouses are family heads
Divorces, separations [2] and annulments [2] of legal marriages and equivalents [2] of common-law unions	If there are many children and custody is shared	If one parent has custody	If the couple is childless.
Deaths of married persons [1] and of persons living in common-law unions [2]		If one spouse in a husband-wife family with children dies, or if a "non-family" married person dies.	If one spouse in a childless husband-wife family, or both spouses in a husband-wife family with or without children, or a married lone-parent family head dies.
Deaths of widows, widowers, or divorcees [1]		If a non-family person dies	If a lone-parent family head dies
Death [2] or leaving home [2] of a lone-parent-family child		If at least another child still lives at home	If a lone child dies or leaves home
Births [1]	If a child is born to a non-family person or to a child at home	If a child is born to a head of a lone-parent family or to a husband and wife	
International migration [1]	If a family immigrates		If a family emigrates
Interprovincial migration [1]	If a family arrives from another province [3]		If a family leaves for another province [3]

[1] Refers to events for which insufficient information is available.
[2] Refers to events for which no statistical information is available.
[3] Applies only at the provincial level.

listed in Table 7.1. Information on marriages, divorces, deaths, immigration, emigration and interprovincial migration (the important family formation/dissolution factors), is either readily available, or can be derived indirectly from administrative data sources. The lack of information on *de facto* separations, annulments, and on children leaving home, however, makes a complete accounting of factors contributing to the formation (or dissolution) of a family, almost impossible. Fortunately, however, the events for which little or no statistical information is available are infrequent, and their anticipated impact is small. More serious is the problem of the unavailability of current data on common-law unions.

The total number of census families in the year t + 1 is estimated using the following equation:

$$F_{t+1} = F_t + [F^m_{(t,\,t+1)} + I_{(t,\,t+1)} + N_{(t,\,t+1)}] -$$

$$[E_{(t,\,t+1)} + \mathcal{V}_{(t,\,t+1)} + \mathfrak{D}^m_{(t,\,t+1)} + \mathfrak{D}^w_{(t,\,t+1)} + \mathfrak{D}^v_{(t,\,t+1)}] + \mathcal{U}_{(t,\,t+1)} \tag{1}$$

where:

F_{t+1} = number of census families in year t + 1;

F_t = number of census families in census year t;

F^m = new families resulting from marriages;

I = families immigrating from another country;

N = net interprovincial migrant families;

E = families emigrating to another country;

$\mathcal{V}$ = dissolution of families due to divorce;

$\mathfrak{D}^m$ = dissolution of families due to deaths of married persons;

$\mathfrak{D}^w$ = dissolution of families due to deaths of widowed persons;

$\mathfrak{D}^v$ = dissolution of families due to deaths of divorced persons;

$\mathcal{U}$ = net change in common-law unions;

(t, t + 1) = time interval from June 1 of year t to May 31 of year t + 1.

The derivation of the above components is addressed in the following sections.

(1) The Contribution of Marriages to Family Formation

For estimation purposes, marriages must be classified by age, marital status and the headship status of both partners immediately preceding marriage. Vital Statistics provides data on marriages, by age and prior marital status, separately for males and females. Vital Statistics does not, however, provide information on headship status. The latter information can be estimated from census data. The combined information on marriages (from Vital Statistics) and headship rates (derived from the last census) is used to estimate the number of newly formed families.

The first step in the process is the estimation of the number of marriages, separately for each of the possible sex and marital status combinations, where at least one spouse was the head of a lone-parent family prior to the current marriage. This is accomplished by applying the headship rate for lone-parent families of a given marital status to the number of marriages for a given sex and marital status, both at a given age. The age, sex and marital status-specific number of marriages for non-family heads is obtained by subtracting marriages of lone-parent heads from the corresponding total number of marriages. This operation is summarized in equations (2), (3) and (4) respectively, and is performed separately for each sex, as follows:

$$M_a^{ms}(h) = M_a^{ms} \cdot H_a^{ms} \tag{2}$$

$$H_a^{ms} = \frac{L_a^{ms}}{P_a^{ms}} \tag{3}$$

$$M_a^{ms}(\bar{h}) = M_a^{ms} - M_a^{ms}(h) \tag{4}$$

where:

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

M_a^{ms} = total number of marriages by marital status and age;

H_a^{ms} = headship rate by marital status and age;

L_a^{ms} = number of persons who are lone-parent family head by marital status and age, enumerated in the last census;

P_a^{ms} = total number of persons by marital status and age enumerated in the last census;

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

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

$$\pi^{ms}(h') = \sum_{ms'} \left[\frac{M^{ms,\,ms'}}{M^{ms}} \cdot \frac{M^{ms'}(h')}{M^{ms'}} \right] \tag{5}$$

probability that a male (female) of marital status ms(ms′) has married a female (male) of marital status ms′(ms)	×	probability that a newlywed female (male) of marital status ms′(ms) is already a family head, h′(h)

where:

$$M^{ms'}(h') = \sum_{a'} (M_{a'}^{ms'} \cdot H_{a'}^{ms'})$$

and,

$$\pi^{ms}(\bar{h}') = 1 - \pi^{ms}(h') \tag{6}$$

where:

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

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

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

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

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

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

Category 1: male only is family head

$$M_a^{ms}(h, \bar{h}') = M_a^{ms}(h) \cdot \pi^{ms}(\bar{h}') \tag{7}$$

Category 2: female only is family head

$$M_a^{ms}(\bar{h}, h') = M_a^{ms}(\bar{h}) \cdot \pi^{ms}(h') \tag{8}$$

Category 3: both male and female are family head

$$M_a^{ms}(h, h') = M_a^{ms}(h) \cdot \pi^{ms}(h') \tag{9}$$

Category 4: neither male nor female is family head

$$M_a^{ms}(\bar{h}, \bar{h}') = M_a^{ms}(\bar{h}) \cdot \pi^{ms}(\bar{h}') \tag{10}$$

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

$$\pi_a^{ms} = \frac{M_a^{ms}(\bar{h}, \bar{h}') - M_a^{ms}(h, h')}{M_a^{ms}} \tag{11}$$

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

$$F^m_{(t,\,t+1)} = \tfrac{1}{2}\left[\sum_{ms}\sum_{a}\left(\pi_a^{ms} \cdot M^{ms}_{a,(t,\,t+1)}\right) + \sum_{ms'}\sum_{a'}\left(\pi_{a'}^{ms'} \cdot M^{ms'}_{a'(t,\,t+1)}\right)\right] \tag{12}$$

(2) The Contribution of Divorce to Family Dissolution

When a couple who have no children living at home undergo a divorce, a family is dissolved. If there are children living at home, however, the effect of divorce on net family formation becomes indeterminate, due to the lack of cross-classified information on child custody by place of residence, as well as the number of children living at home. The estimated number of family dissolutions resulting from divorce is, therefore, assumed to be equal to the number of divorces[3] involving couples with no dependent children. (Dependent children are defined as all children under the age of 16, and all children aged 16+ who are financially dependent.)

(3) The Contribution of Death to Family Dissolution

The families at risk of being dissolved through death can be: two-parent families with or without children, and lone-parent families whose head is either single, widowed, separated or divorced. No information is available on, nor has any attempt been made to estimate, family dissolution through either of the following categories: (1) death of an only child in lone-parent families, or: (2) death of either children, or both parents, in the case of two-parent families. These cases are rare and have been ignored here. The following sets out the procedures used for estimating the dissolution of family types.

Couples Without Children and Lone-Parent Families Whose Head is a Separated Person

Since Vital Statistics death data do not differentiate deaths of separated persons from those of other married persons, the separate estimation of the number of family dissolutions due to deaths in husband-wife families, and in lone-parent families (where the head of the family is separated) cannot be accomplished. An estimation of the total number of such family dissolutions through death is, therefore, effected by multiplying the number of deaths of married persons, by age group, by age-specific probabilities of family dissolutions due to death, as shown below:

$$\mathfrak{D}^{m}_{(t,\, t+1)} = \sum_{a} \left(D^{m}_{a,\,(t,\, t+1)} \cdot \pi^{m}_{a}\right) \tag{13}$$

where: $\mathfrak{D}^{m}_{(t,\, t+1)}$ = number of families dissolved through deaths of married persons during the period t to t + 1;

$D^{m}_{a,\,(t,\, t+1)}$ = number of deaths of married persons by age group during the period t to t + 1;

π^{m}_{a} = age-specific probabilities of family dissolution due to deaths of married persons.

---

[3] Of the divorce rulings issued between June 1, 1982 and May 31, 1983, 45 % involved couples with no dependent children, 22 % involved couples with one dependent child and 33 % involved couples with more than one dependent child.

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

$$\pi_a^m = \frac{P_a^m(c) + L_a^m}{P_a^m} \tag{14}$$

where: π_a^m = probability of family dissolution due to deaths of married persons, by age group;

$P_a^m(c)$ = number of couples without children, by age group, as enumerated in the last census;

L_a^m = number of lone parents who are separated, by age group, as enumerated in the last census;

P_a^m = total number of married persons, by age group, as enumerated in the last census.

Dissolution of Families Whose Head is Widowed or Divorced

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

$$\mathfrak{D}^w_{(t,\,t+1)} = \sum_a (D^w_{a,\,(t,\,t+1)} \cdot H^w_a) \tag{15}$$

$$\mathfrak{D}^v_{(t,\,t+1)} = \sum_a (D^v_{a,\,(t,\,t+1)} \cdot H^v_a) \tag{16}$$

where,

$$H^w_a = \frac{L^w_a}{\hat{W}_a} \tag{17}$$

and

$$H^v_a = \frac{L^v_a}{\hat{V}_a} \tag{18}$$

$(t, t+1)$ = time interval from June 1 of year t to May 31 of year t + 1;

$\mathfrak{D}^{w}_{(t, t+1)}$ and $\mathfrak{D}^{v}_{(t, t+1)}$ = number of families dissolved through deaths of widowed and divorced persons, respectively;

$D^{w}_{a, (t, t+1)}$ and $D^{v}_{a, (t, t+1)}$ = number of deaths of widowed and of divorced persons, by age group, respectively;

H^{w}_{a} and H^{v}_{a} = headship rate of widowed population and of the divorced population, respectively, by age group, calculated from the last census data;

L^{w}_{a} and L^{v}_{a} = number of lone-parents who are widowed and divorced, respectively, by age group, in the last census;

$\hat{W}_{a}$ and $\hat{V}_{a}$ = total number of widowed and divorced persons, respectively, by age group, enumerated in the last census.

The total number of families dissolved by the death of widowed and divorced persons is obtained by summing the results of equations 15 and 16.

(4) Immigration and Emigration of Families

The determination of the number of new families resulting from immigration requires information on the number of family units (according to the definition of census family) who immigrate, and on the number of family reunifications resulting from the arrival of persons who already have a parent, child or spouse in Canada. As neither is available, certain compensatory assumptions have to be made. It is reasonable to assume that the number of immigrant or emigrant families will be closely approximated by the total number of married immigrant or emigrant females. It is posited, therefore, that for each married immigrant or emigrant female, there is one immigrant or emigrant family.

In recent years approximately 50 % of all immigrants were female. Furthermore, almost 50 % of these women were married when they arrived in Canada. This means that there is roughly one family for every four immigrants.

(5) Interprovincial Migration of Families

Interprovincial family migrations are calculated from Family Allowance files (see Chapter IV for a more complete description of these files). Each month, Health and Welfare Canada tabulates the number of Family Allowance accounts transferred from one province to another, by province of origin and destination. The definition of family used for family allowance purposes does not correspond exactly to the broader definition of a census family, since only families with children under 18 years of age are entitled to family allowances. The number of families whose migration has been recorded in the family allowance master file is, therefore, inflated by using the ratio of the number of families enumerated in the last census to the number of family accounts in the family allowance master file, for June of the last census year. These ratios, for June 1981, are shown in Table 7.2.

(6) Contribution of Common-Law Families to Family Formation

For each province, age and sex-specific ratios measuring annual change in the number of persons living in common-law relationships, are calculated using data from the two latest censuses in conjunction with population estimates by age group and sex. These annual ratios are applied to the corresponding postcensal estimates of population as follows:

$$\Delta u_{a,(t,t+1)} = \left[\frac{1}{5} \left(\frac{U_{a,t} - U_{a,t-5}}{(P_{a,t} + P_{a,t-5}) \div 2} \right) \right] \cdot \hat{P}_{a,(t+1)} \tag{19}$$

where, for each sex:

$\Delta u_{a,(t,t+1)}$ = annual change in the number of persons living in common-law relationships, by 5-year age group, during a one-year period;

U_a = number of persons in age group "a" living in common-law relationships as enumerated in the most recent censuses;

P_a = enumerated population in age group "a";

$\hat{P}_a$ = estimated population in age group "a".

The net annual change in the number of persons living in common-law relationships ($\Delta u_{(t,t+1)}$) is then translated into changes in the number of common-law families, by averaging the number of males and females living in common-law relationships, as follows:

$$\mathcal{U}_{(t,t+1)} = \tfrac{1}{2} \left[\sum_{s} \sum_{a} \Delta u_{a,(t,t+1)} \right] \tag{20}$$

where: a = 5-year age groups (from 15-19 to 60-64 and 65+);
s = sex.

TABLE 7.2. Number and Ratio of Families in Census to Number of Family Allowance Accounts, 1981

Geographic Area	Census Families (1)	Families in Family Allowance File (2)	Ratio (3) = (1) ÷ (2)
Canada	**6,324,323**	**3,602,004**	**1.76**
Newfoundland	135,138	96,453	1.40
Prince Edward Island	30,217	18,501	1.63
Nova Scotia	216,182	128,846	1.68
New Brunswick	176,550	110,078	1.60
Quebec	1,671,395	958,404	1.74
Ontario	2,278,783	1,264,626	1.80
Manitoba	262,178	147,346	1.78
Saskatchewan	245,650	141,958	1.73
Alberta	565,507	332,877	1.70
British Columbia	727,569	391,079	1.86
Yukon	5,675	3,817	1.49
Northwest Territories	9,479	8,019	1.18

Source: The 1981 Census figures adjusted to June 1, 1981.
Health and Welfare Canada, *Monthly Statistics, Income Security Programmes, Family Allowances*, June 1981, Table 2 (Regular Account only).

Family Characteristics

The following family characteristics are estimated:

(a) Size of family (2, 3, 4, 5 and 6+ members) and total number of persons in family;

(b) Type of family, as husband-wife or lone-parent;

(c) Broad age group of children;

(d) Broad age group of husband and wife;

(e) Broad age group and sex of lone-parent.

The number of families having the above characteristics is estimated using a linear extrapolation of trends based on data from the two most recent censuses. The following four operations are involved in this procedure:

(1) Extrapolation of proportions of families by characteristics for each province;

(2) Adjustment for extrapolated proportions less than zero;

(3) Estimating the number of families by characteristics;

(4) Estimating the number of persons in census families.

(1) Extrapolation of Proportions

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

$$\Phi^k_{t+i} = \Phi^k_t + \left[i/5 \left(\Phi^k_t - \Phi^k_{t-5} \right) \right] \tag{21}$$

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

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

(2) Adjustment for Negative Proportions

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

$$\Phi'^k_{t+i} = \Phi''^k_{t+i} \cdot \left[\frac{1}{\sum \Phi''^k_{t+i}} \right] \tag{22}$$

where: Φ'^k_{t+i} = adjusted cell value after removing negative proportions from the table.

Φ''^k_{t+i} = unadjusted cell values after removing negative proportions. These values are zero where Φ^k_{t+i} is negative, and equal to Φ^k_{t+i} otherwise (i.e., where $\Phi^k_{t+i} \geq 0$).

(3) Estimated Number of Census Families by Characteristics

The estimated number of families for the year t + i with characteristic "k" is calculated by multiplying the total number of families F_{t+i} (estimated by the component method described in the first section of this chapter) by the adjusted proportions (Φ'^k_{t+i}), as below:

$$F^k_{t+i} = F_{t+i} \cdot \Phi'^k_{t+i} \tag{23}$$

(4) Estimated Number of Persons in Census Families

Estimating the total number of persons in census families requires several steps, due to the existence of the open-ended family size category, "six-plus". First, the proportion of all persons in census families at the time of the last two censuses (Φ_t and Φ_{t-5}), for any province, is calculated from the appropriate census data. Taking the most recent census as the base, the proportion of all persons in census families for the estimate year (Φ_{t+i}) again by province, is calculated by linear extrapolation as follows:

$$\Phi_{t+i} = \Phi_t + (i/5\ (\Phi_t - \Phi_{t-5})) \tag{24}$$

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

$$N_{t+i} = \hat{P}_{t+i} \cdot \Phi_{t+i} \tag{25}$$

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

$$\Phi^{6+}_{t+i} = \Phi^{6+}_t + i/5\ [\Phi^{6+}_t - \Phi^{6+}_{t-5}] \tag{26}$$

Next, the estimated *number* of persons in census families of at least size 6 (N^{6+}_{t+i}) is calculated by applying the result of equation 26 to that of equation 25. Specifically:

$$N^{6+}_{t+i} = N_{t+i} \cdot \Phi^{6+}_{t+i} \tag{27}$$

Finally, having derived the number of families by sizes 2 through 5 from equation (23) and the estimate of the number of persons in families of size 6+, the adjusted estimate of the total number of persons living in census families N'_{t+i} can be obtained by using the following equation:

$$N'_{t+i} = \left[\sum_{n=2}^{5} [n(F^n_{t+i})] + N^{6+}_{t+i} \right] \tag{28}$$

Evaluation of Estimates

One method of evaluating family estimates is to compare the estimates for the census year to the census counts. Table 7.3 presents a comparative measure, the error of closure, for 1976 and 1981. Two important features are revealed. First, there is an apparent understatement of the estimated number of families at both points in time. Second, and more importantly, the absolute size of this understatement almost doubled when the 1981 and 1976 Censuses were compared. For Canada, the error of closure increased from −1.13 % in 1976, to −2.10 % in 1981.

Four sources of error and bias contribute to the error of closure : (1) common-law unions; (2) census undercoverage; (3) the universes covered, and; (4) unrecorded events.

Common-Law Unions

While persons in common-law unions are counted as being "married" in the censuses, they are not subject to Vital Statistics registration. Data from the 1981 Census and the 1984 Family History Survey indicate that, since the mid-1970's, there has been a substantial rise in common-law living arrangements particularly among individuals in their twenties. About two-thirds of the difference between the estimated and enumerated number of families in 1981 can be attributed to the differential treatment of common-law unions between Census and Vital Statistics.

TABLE 7.3. Error of Closure of Census Families: Canada and Provinces, June 1, 1976 and 1981

Geographic Area	1976			1981		
	Census	Error of Closure		Census[2]	Error of Closure	
		Number	Percent[1]		Number	Percent[1]
Canada[3]	**5,714,724**	**−64,636**	**−1.13**	**6,309,169**	**−132,231**	**−2.10**
Newfoundland	124,650	−682	−0.55	135,100	403	0.30
Prince Edward Island	27,560	193	0.70	30,200	69	0.23
Nova Scotia	200,480	−1,786	−0.89	216,200	−3,961	−1.83
New Brunswick	162,030	852	0.53	176,600	−1,683	−0.95
Quebec	1,540,400	−47,628	−3.09	1,671,400	−41,998	−2.51
Ontario	2,104,540	−2,816	−0.13	2,278,800	−26,815	−1.18
Manitoba	251,970	−259	−0.10	262,200	−5,629	−2.15
Saskatchewan	225,685	614	0.27	245,700	−3,095	−1.26
Alberta	448,765	1,614	0.36	565,500	−20,582	−3.64
British Columbia	628,445	−14,738	−2.35	727,600	−28,940	−3.98
Mean Absolute Error			0.90			1.80

[1] Percent error of closure $= \frac{\text{Estimates - Census}}{\text{Census}} \times 100$

[2] The June 3, 1981 Census figures were adjusted to refer to June 1, 1981.
[3] Excluding Yukon and Northwest Territories.
Source: Statistics Canada, Demography Division.

Census Undercoverage

No estimates of census undercoverage of families are available. Some inferences can, however, be drawn from the estimates of census undercoverage of private households, a concept close to that of census family. As shown in Table 7.4, the undercoverage was higher in 1976 (1.97 %) than in 1981 (1.71 %), suggesting that the number of households and, by implication, the number of families in 1976, was slightly understated relative to the 1981 Census, thus contributing to the higher error of closure in 1981.

Family Universes

Family data from the census is not compatible with family data available from the Vital Statistics registration system, due to the differential treatment of collective households[4] and the events affecting their formation or dissolution. These families are not included in census family tabulations, while the vital events experienced by the members of families or non-family persons in collective households are registered by Vital Statistics. This results in a discrepancy between families as identified by the census, and families as estimated on the basis of family-related events captured by Vital Statistics. The magnitude of error associated with this discrepancy is, however, unknown.

TABLE 7.4. Private Household Undercoverage Rate and Standard Error: Canada[1] and Regions, 1971, 1976 and 1981

Geographic Area	1971		1976		1981	
	Estimated Rate	Standard Error	Estimated Rate	Standard Error	Estimated Rate	Standard Error
Canada[1]	**1.46**	**0.09**	**1.97**	**0.11**	**1.71**	**0.13**
Atlantic	1.10	0.28	1.80	0.30	1.85	0.28
Quebec	1.60	0.19	2.31	0.24	1.72	0.29
Ontario	1.08	0.12	1.52	0.25	1.39	0.23
Prairies	1.33	0.19	1.34	0.16	1.81	0.27
British Columbia	2.81	0.39	3.61	0.39	2.40	0.38

[1] Excluding Yukon and Northwest Territories.

Source: G.I. Brackstone and J.F. Gosselin, *1971 Census Evaluation Programme, MP1 - 1971 Reverse Record Check*, Statistics Canada, Ottawa, 1973.
Quality of Data, 1976 Census of Canada, Catalogue No. 99-840, Statistics Canada, Ottawa, March 1980, Table 5.
1981 Census of Canada. *Data Quality - Total Population*. Catalogue No. 99-904, Statistics Canada, Ottawa (forthcoming).

[4] It is estimated that in 1981, 400,000 persons (1.7 % of the total population) were living in collective households. A collective household, as defined in the 1981 Census, refers to "a person or group of persons who occupy a collective dwelling and do not have a usual place of residence elsewhere in Canada." (Statistics Canada, *1981 Census Dictionary,* Catalogue 99-901, p. 72). Collective dwelling refers to "a dwelling of a commercial, institutional or communal nature." (Statistics Canada, ibid., p. 82).

Unrecorded Events

Unrecorded events, such as: (1) the death or departure from home of the only child of a lone-parent family; (2) separations, annulments of marriages, and; (3) births of children to non-family persons, are all possible sources of error. Some have a positive, while others have a negative effect on family formation. The net effect, however, is probably small.

From the above review of possible sources of error of closure, it is fair to conclude that the inclusion of common-law unions in the censuses, and their exclusion from Vital Statistics registrations, is the major source of discrepancy.

APPENDIX I: GLOSSARY OF PRINCIPAL TERMS

Base Population

The population at the beginning of a period used as a reference or starting point for the estimation process. A base population can be either a **population estimate** or the **enumerated population** (see also **Component Method** and **Regression-nested**).

Birth Cohort

Refers, in this publication, to all persons born within any 12-month period ending May 31st.

Census Division (CD)

A general term applying to counties, regional districts, regional municipalities and five other types of geographic areas made up of groups of census subdivisions. In Newfoundland, Manitoba, Saskatchewan and Alberta, the term describes areas that have been created by Statistics Canada in cooperation with the provinces as an equivalent for counties. (1981 Census Dictionary, Catalogue No. 99-901)

Census Family

A census family is a husband and wife with or without never-married children (regardless of their age), or either parent with one or more **never-married** children living in the same dwelling. Persons living in **common-law unions** are considered to be married and to constitute a **husband-wife family** (see also **Family Structure** and **Lone-Parent Family**).

Census Metropolitan Area (CMA)

The main labour market area of an urbanized core (or continuously built-up area) having 100,000 or more population. They contain whole municipalities (or **census subdivisions**).

Census Subdivision

A general term applying to municipalities, Indian Reserves, Indian Settlements and unorganized territories.

In Newfoundland, Nova Scotia and British Columbia, the term also describes geostatistical areas that have been created by Statistics Canada in cooperation with the provinces as an equivalent for municipalities. (1981 Census Dictionary, Catalogue No. 99-901)

Censal Year

The censal year refers to the period beginning June 1 of any year (not necessarily a year in which there is a census) and ending May 31 of the following year.

Census Year

The calendar year in which a census is taken. In Canada, a census is taken every five years, the most recent in 1986.

Children in Family

Children in the family are defined as **never-married** sons and daughters (including adopted children and stepchildren), regardless of their age,living in the same dwelling as their parents. **Ever-married** sons and daughters of any age are not considered members of their parents' family even if they live in the same dwelling (see also **Census Family, Family Structure** and **Lone-Parent Family**).

Common-Law Union

A male and female, living together as husband and wife, without the sanction of a legal marriage.

Component Method

A method of generating **population estimates** which uses the **components of demographic change** and a **base population** as the input (see also **Regression-nested**).

Components of Demographic Change

The factors responsible for **population growth** and variations in the demographic composition of populations between two points in time. These include: births; deaths; marriages; divorces, and other **vital events,** as well as **migration** in its various forms.

Emigration (see under **Migration)**

Enumerated Population

The population of an area according to an official census (see also **Population Estimate**).

Error of Closure

Difference between the **population** according to a Census and a **postcensal estimate** for the same date.

Error Term

The error term (ε) is added to the equation in a **regression model** in order to account for all the **variance** in the dependent variable. Unless a perfect relationship exists between the dependent and independent variables (which is almost impossible) there will be unexplained variance, defined as the difference between the predicted value in the regression, and the actual (or observed) value. This error is also often referred to as "stochastic error", "residual error", and/or "variance not explained by regression".

Estimate (See **Population estimate)**

Ever-Married

A person is considered ever-married if he/she is currently married, separated, widowed or divorced (see also **Never-Married**).

Family Structure

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

Homogeneity of Variance of the Error Term

Refers to the assumption that the **error term** in a **regression** equation has constant **variance**.

Husband-Wife Family

A husband-wife family consists of a legally married male and female (with or without children), or two persons living in a **common-law union** (with or without children) (see also **Census Family, Family Structure** and **Lone-Parent Family**).

Immigration (see under **Migration)**

Landed Immigrant

An international migrant who has been granted lawful permission to enter Canada to establish permanent residence.

Lone-parent Family

A lone-parent family consists of a father/mother, with one or more **never-married** children living in the same dwelling (see also **Census Family, Family Structure** and **Husband-wife Family**).

Migration

The movement of members of a **population** involving a change in permanent residence.

Emigration

Emigration represents departures from Canada to another country, involving a change in usual place of residence (see also **Immigration** and **International Migration**).

Immigration

Immigration represents entries into Canada of **landed immigrants** from another country, involving a change in usual place of residence (see also **Emigration** and **International Migration**).

Internal Migration

Movement between geographical units within Canada involving a change in usual place of residence (see also **Interprovincial Migration** and **Subprovincial Migration**).

International Migration

Movement in either direction between Canada and other countries involving a change in usual place of residence (see also **Emigration** and **Immigration**).

Interprovincial Migration

Interprovincial migration represents movements from one province to another involving a change in usual place of residence (see also **Subprovincial Migration** and **Internal Migration**).

Subprovincial Migration

Movement between geographical units within a province involving a change in usual place of residence (see **Internal Migration** and **Interprovincial Migration**).

Multicollinearity

Multicollinearity exists when some or all of the independent variables in a **regression model** are highly correlated among themselves.

Never-married

A person who is not now, and has never been, married (see also **Ever-married**).

Paired t — test

A statistical test that determines if the difference between two sets of paired observations is statistically significant.

Population

Estimated population **(population estimate)** and population according to the census **(enumerated population)** are both defined as being "the total number of persons whose usual place of residence ... was somewhere in Canada, including Canadian government employees stationed abroad and their families, members of Canadian Armed Forces stationed abroad and their families, and crews of Canadian merchant vessels. Not included are government representatives of other countries and their families attached to the legation, embassy, or other diplomatic body of that country, members of the Armed Forces of other countries stationed in Canada and members of their families who are not citizens of Canada, students attending school in Canada whose usual residence is outside Canada, (workers from another country in Canada on Employment Visas and their families) and residents of another country visiting in Canada temporarily" (Dictionary of the 1971 Census terms, Catalogue No. 12-540, page 28).

Population Estimate

A non-census accounting of the current size of a population and/or its spatial distribution and/or demographic characteristics (see also **Population Projections, Postcensal Estimate, Intercensal Estimate**).

Intercensal Estimate

Population estimate for intercensal years derived by using **postcensal estimates** and population counts from the last and preceding censuses (see also **Population Estimate** and **Postcensal Estimate**).

Postcensal Estimate (Preliminary, Updated and Final)

Population estimate produced by using data from the most recent census and estimates of the **components of demographic change** since that census (see also **Population Estimate** and **Intercensal Estimate**).

Population Growth

Change in population size from one date to another. Growth can be positive (increase) or negative (decrease).

Population Projections

As distinct from **population estimates** , a projection is an attempt to quantify what the size, spatial distribution and/or demographic characteristics of a population will be in the future, based on assumptions about future demographic trends. Population estimates are often used as the **base population** for projections.

Regression Model

A statistical equation that derives a relationship between a dependent variable (i.e. change in population) and a set of independent variables (i.e. change in family allowance recipients or health care reference population). The equation derived from historical data is then used with current data on the independent variables to obtain **population estimates** (see also **Symptomatic Indicator**).

Regression-nested

A method of producing **population estimates** which involves combining the change in population from one time to another, produced by the regression method, with the previous year's population estimate obtained by the component method.

Stochastic Error (see **Error Term**).

Symptomatic Indicator

Independent variable used in a **regression model** . The symptomatic indicators for **population estimates** are taken from administrative files such as family allowances, health care and hydro accounts.

Variance

The sum of the squared deviation of each observation from the arithmetic mean of all observations (see also **Error Term, Regression Model** and **Homogeneity of Variance of the Error Term**).

Vital Events

Any event such as birth, death, marriage, divorce for which there exists a legal requirement to file a notification with the Provincial or Territorial Registrar's Office.

APPENDIX II: SYMBOLS AND THEIR MEANING

Population and Demographic Events

B = Number of births

D = Number of deaths

D^m = Number of deaths of married persons

E = Number of emigrants

f = separation factor

f = Estimation factor for migration

I = Number of immigrants

M = Number of marriages

$\hat{\mathrm{M}}$ = Estimates of married population

M = Number of out-migrants

$\hat{M}$ = Estimates of out-migrants

N = Number of net internal migrants

P = Population counts in Census

$\hat{\mathrm{P}}$ = Population estimates

$\mathcal{P}$ = Intercensal population estimates

$\hat{\mathrm{S}}$ = Estimates of single population

V = Number of divorces

$\hat{\mathrm{V}}$ = Estimates of divorced population

W = Number of persons who became a widow(er)

$\hat{\mathrm{W}}$ = Estimates of widowed population

Family

$\mathfrak{D}$ = Dissolution of families due to deaths

E = Emigrating families to another country

F = Number of census families

$\hat{F}$ = Estimated number of census families

F^m = New families resulting from marriage

H = Headship rate

h = Male head of a lone-parent family

h′ = Female head of a lone-parent family

I = Immigrating families from another country

k = Characteristic k

L = Lone-parent families

ms = Marital status

N = Number of net interprovincial migrant families

N = Total number of persons in families

N′ = Adjusted total number of persons in census families

n = family size (used as a superscript)

U = Number of persons living in common-law families

$\mathfrak{U}$ = Change in common-law families

$\mathfrak{V}$ = Dissolution of families due to divorce

Note: Symbols E, I and N are used for migration of either individuals or families.

Characteristics (superscript or subscript)

a = Age

a′ = Cohort age/Age of spouse

m = Married

ms = Marital status

s = Single (never married)

v = Divorced

w = Widowed

Time Dimensions

i = Unit of time

t = Year

Geographic Area

j = A given area or origin of migration (province or smaller unit)

k = Destination of migration (province or smaller unit)

Others

Φ = Proportion

π = Probability

$\triangle$ = Difference

c = Canada

cd = Census division

cm = Census metropolitan area

ε = Error of closure

f = Separation factor

UK = United Kingdom

US = United States

DATE DUE

DEC - 5 1999			
AUG AOUT - 8 2000			
AUG AOUT 2 5 2000			
SEP 1 1 2000			
JUN - 1 2001			
FEB 0 7 2003			
SEP 1 2005			
DEC 0 7 2005			