Survey Methodology, December 1989
Vol. 15, No. 2, pp. 261–270
Statistics Canada

The Use of Administrative Records for Estimating Population in Canada[1]

RAVI B.P. VERMA and RONALD RABY[2]

ABSTRACT

This paper examines the adequacy of estimates of emigrants from Canada and interprovincial migration data from the Family Allowance files and Revenue Canada tax files. The application of these data files in estimating total population for Canada, provinces and territories, was evaluated with reference to the 1986 Census counts. It was found that these two administrative files provided consistent and reasonably accurate series of data on emigration and interprovincial migration from 1981 to 1986. Consequently, the population estimates were fairly accurate. The estimate of emigrants derived from the Family Allowance file could be improved by using the ratio of adult to child emigrant rates computed from Employment and Immigration Canada's immigration file.

KEY WORDS: Interprovincial migration; Emigration; Population estimates; Census counts; Accuracy.

1. INTRODUCTION

The national Census, conducted every five years since 1951, provides a wide range of demographic data on the Canadian population. However, unlike some other industrialized countries, Canada does not have a continuous population registration to derive basic demographic data and track the movement of people over different geographic areas for non-census years. To fill this gap, since the 1940s Statistics Canada has developed a program of population and family estimates. For example, population estimates for Canada, provinces and territories, census divisions, and census metropolitan areas are produced using the latest census counts and several administrative data sources, including: Revenue Canada tax files and Family Allowance files for migration; Vital Statistics registration for births and deaths; and Immigrant Visa and Record of Landing Registration for immigration.

The strengths and weaknesses of these administrative files for estimating population and migration compared with 1981 Census data have been discussed elsewhere. (Statistics Canada 1987; Verma and Parent 1985; Norris, Britton and Verma 1982). In this paper, the accuracy of estimates of the components of population change for provinces and territories using the Family Allowance and Revenue Canada data sources will be evaluated by comparison with the 1986 Census counts. This evaluation will compare 1971, 1976 and 1981 data.

The paper is presented in the following sections: data sources and the methods of estimation; results of the evaluation; and conclusions and discussion.

2. DATA SOURCES AND THE METHODS OF ESTIMATION

This section describes the procedures for estimating total population, interprovincial migration, and emigration.

---

[1] Revised version of a paper presented at Statistics Canada Symposium on Statistical Uses of Administrative Data, November 1987.

[2] Ravi B.P. Verma and Ronald Raby, Demography Division, Statistics Canada, 4-A Jean Talon Building, Ottawa, Ontario, K1A 0T6.

2.1 Total Population

Quarterly and annual estimates of the total population of Canada and the provinces and territories, and annual totals for census divisions and census metropolitan areas, are produced by the component method. At the national level, the number of births and immigrants are added to, and the number of deaths and emigrants subtracted from, the base population (taken from the latest Census of Canada). By province and for smaller areas, estimates of internal migration are also taken into account.

The component method is expressed as follows:

$$\hat{P}(t+i) = P(t) + [B(t,t+i) - D(t,t+i) + I(t,t+i) - E(t,t+i)] + N(t,t+i). \quad (1)$$

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 $t+i$

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

N = number of net interprovincial immigrants between time t and $t+i$

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

2.2 Interprovincial Migration

Two administrative files are used to produce annual and quarterly estimates of interprovincial migration. Preliminary estimates are derived from Family Allowance files, while final figures are estimated from Revenue Canada income tax files.

2.2.1 Preliminary Estimates

The number of adult migrants is estimated using child migration figures derived from Family Allowance files, and ratios of adult out-migration rates to child out-migration rates ($f_{j,k}$) based on the most recent Revenue Canada tax file (calculated for 1 or 2 years before the reference date). Recipients of Family Allowance cheques must notify the Department of Health and Welfare of changes in address. These changes are compiled monthly for both province of origin and destination, by size of family (the number of children per family receiving the allowance). Coverage of the population by Family Allowance is comparable to that of the census (Statistics Canada 1987, p. 46). Estimates of the number of interprovincial out-migrants for all age groups are calculated as follows:

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

$$f_{(j,k)} = \frac{M'_{(j,k),18+}}{\hat{P}_{j,18+}} \div \frac{M'_{(j,k),0\text{-}17}}{\hat{P}_{j,0\text{-}17}} \quad (3)$$

$$\hat{M}_{(j,k),0+} = \hat{M}_{(j,k),18+} + M_{(j,k),0\text{-}17} \tag{4}$$

where:

$\hat{M}_{(j,k),0+}$ = estimated total number of persons out-migrating from province j to province k

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

$M'_{(j,k),18+}$ = number of adult out-migrants from province j to province k derived from Revenue Canada tax files

$M'_{(j,k),0\text{-}17}$ = number of child out-migrants (aged 0-17) from province j to province k derived from Revenue Canada tax files

$M_{(j,k),0\text{-}17}$ = number of child out-migrants from province j to province k, based on Family Allowance files

$P_{j,18+}$ = estimated number of adults in province j, the difference between the total population estimates and estimates of the child population based on Family Allowance files

$P_{j,0\text{-}17}$ = total number of children receiving Family Allowance payments in province j

$f_{(j,k)}$ = estimation factor for adult migrants from province of origin j to province of destination k, based on estimates of migration from Revenue Canada tax files

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

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

2.2.2 Final Estimates

Revenue Canada tax files are used to produce final estimates of interprovincial migrants. All individuals receiving an annual income above a specified minimum are required to file an income tax return by the end of April of each year. Migrant tax filers are identified by comparing area of residence from two consecutive tax returns. Information on the number and ages of dependents is imputed from the total amount of personal exemptions claimed by filers. An adjustment is made for segments of the population not covered by the Revenue Canada system; this includes people who neither file an income tax return nor appear as dependents in another filer's return (Norris and Standish 1983; Statistics Canada 1987).

2.3 Emigration

In Canada no system exists for recording emigrants; hence, their numbers must be estimated. Revenue Canada income tax files with an "out-of-Canada" address one year and an "in-Canada" address for the previous year are used to identify emigrants. The emigrant status of children under 17 years of age is determined from change of address notifications from Family Allowance recipients. By combining information from these two administrative files, both preliminary and final estimates of emigrants are generated. The estimation procedures are similar to those used to estimate preliminary interprovincial migration:

$$\hat{E}_j = \left[\frac{E_{j,0\text{-}17}}{P_{j,0\text{-}17}} \cdot f_c \cdot P_{j,18+}\right] + E_{j,0\text{-}17} \tag{5}$$

$$f_c = \frac{E'_{c,18+}}{\hat{P}_{c,18+}} \div \frac{E'_{c,0\text{-}17}}{\hat{P}_{c,0\text{-}17}} \tag{6}$$

$$\hat{E}_c = \sum_{j=1}^{12} \left[\hat{E}_j\right] \tag{7}$$

where:

$\hat{E}_j$ = estimated annual number of emigrants from province j

$\hat{E}_c$ = estimated annual number of emigrants from Canada

$E_{j,0\text{-}17}$ = number of emigrants from province j aged 0 to 17 who were eligible for Family Allowance

$P_{j,0\text{-}17}$ = number of children in province j who are eligible for Family Allowance

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

f_c = annual adjustment factor for estimating adult emigration from Canada, based on Revenue Canada tax files.

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

$\hat{P}_{c,18+}$ and $\hat{P}_{c,0\text{-}17}$ = estimated June 1st population of adults and children for Canada, based on the component method.

The method of estimating the number of emigrants was modified in March 1989, affecting estimates after 1986. The new method combines counts by age of emigrants from Canada to the United States (from the U.S. Department of Justice, Immigration and Naturalization Service), and estimates of the numbers of emigrants from Canada to countries other than the U.S. based on Family Allowance files and an f_c factor calculated from immigration files (see Raby, Martel and Cartier 1989).

3. EVALUATION OF ESTIMATES OF THE COMPONENTS OF POPULATION CHANGE

Each component of population change (births, deaths, immigrants, emigrants and interprovincial migrants) may contain a degree of bias and error. However, the data on births, deaths and immigration can be regarded as more accurate than the estimates of emigrants and interprovincial migrants. In 1982, the methods of estimating emigrants and internal migration were thoroughly updated (see Statistics Canada 1987). These revised methods are evaluated below.

Table 1

Estimates of Emigrants by Different Methods, Canada, 1976-1981 and 1981-1986

Method	1976-81	1981-86
Residual*		
(a) Unadjusted	277,558	476,373
(b) Adjusted for Undercoverage	196,955[1]	134,857[1]
(c) Adjusted for Net Undercoverage	194,155[2]	218,148[2]
Revenue Canada Tax File	207,420	165,272
Family Allowance Method	278,624	235,481
Reverse Record Check	296,724	288,376

*Residual Method:
Emigrants = ([Births − Deaths] + [Immigrants]) − Intercensal growth of population between time t and $t + 5$.

[1] The undercoverage rates were 2.04% for the 1976 Census, 2.01% for the 1981 Census, and 3.21% for the 1986 Census.
[2] The 1976, 1981 and 1986 Census net undercoverage rates were 1.53%, 1.51% and 2.40% respectively. They are estimated using the U.S. experience of overcoverage which is 25% of the undercoverage rate.
Source: Demography Division, Statistics Canada.

3.1 Emigration Data

Table 1 presents estimates of emigrants from Canada by using different methods and data sources for 1976-1981 and 1981-1986. For 1981-1986, the estimate using the residual method is considerably higher than the estimate based on the Family Allowance file. The residual method subtracts the population growth between 1981 and 1986, unadjusted for census undercoverage, from natural increase and immigration. Since births, deaths and immigration data are assumed to be accurate, the higher estimate by the residual method can be attributed to the difference in undercoverage rates for 1981 and 1986. After adjusting the 1981 and 1986 Census counts for undercoverage (2.01% and 3.21% respectively), the estimate by the residual method was found to be 134,857. This figure is lower than estimates obtained using both the Family Allowance file (235,481) and the Revenue Canada tax file (165,272).

This low estimate may result from different rates of overcoverage in the 1981 and 1986 Censuses. No estimate of overcoverage is calculated in the Reverse Record Check study, but the rate can be assumed to be similar to the U.S. Census rate which is 25% of the undercoverage rate. After adjusting the 1981 and 1986 Census counts for net coverage rates of 1.51% and 2.40% respectively, the residual estimate (218,148) was close to the Family Allowance-based estimate (235,481).

For 1976-1981, the estimating methods do not produce similar results. The number of emigrants estimated by the residual method adjusted for net undercoverage was 194,155, which is close to the estimate based on Revenue Canada tax files (207,420), but considerably lower than the Family Allowance method estimate (278,624) or the Reverse Record Check estimate (296,724).

One possible source of error in the current method is the f_c factors, which are adult-child emigrant ratios, estimating the number of emigrants aged 18+ from 1981-1986. These ratios were obtained from the emigration data provided by the Revenue Canada tax files.

Table 2 shows f_c values derived from different data sources. The f_c factors from the Revenue Canada tax files are less than unity and higher than unity from the three other data sources: interprovincial migration data from income tax files, immigration files, and data on Canadian emigrants to the United States. The estimates of emigrants from these sources are also higher than the Revenue Canada-based estimate.

Table 2

Estimates of Emigrants by Family Allowance Method Using Different Values of f_c (Adult-Child Emigrant Ratios), 1981-1986

Data Source of f_c	Value of f_c Factor					Number of Emigrants
	1981-82	1982-83	1983-84	1984-85	1985-86	
1. Revenue Canada Tax Files	0.8698	0.8768	0.9052	0.8592	0.8592	235,481
2. Interprovincial Migration Data from Income Tax Files	1.0760	1.1000	1.0664	1.0290	1.0029	265,816
3. EIC Immigration Data	1.0801	1.0926	1.1723	1.1254	1.0694	275,762
4. Canadian Emigrants to the U.S.A.	1.2300	1.2774	1.3196	1.3745	1.4232	316,268

Source: Demography Division, Statistics Canada.

Each f_c factor source shows annual variations. The f_c factors for Canadians emigrating to the United States are particularly high, indicating that 23% to 42% more adults emigrated to the U.S. than did children. This is not surprising, as the southern American states have always been attractive to retirees. Hence the f_c factor based on U.S. data may not be suitable for estimating Canadian emigrants to countries other than the U.S.

Similarly, the f_c factors for interprovincial migration, based on the income tax file, suggest that adult migrants have exceeded child migrants by up to 10% from 1981 to 1986. However, the adult migrant group likely contains a high proportion of younger adults, who tend to move more often between provinces than other age groups. Hence this data source is also very specific and thus not suitable for computing the overall f_c factor.

According to some authors (Beaujot and Rappak 1988), emigrant and immigrant flow data are associated, making it possible to compute an f_c factor from the Emloyment and Immigration Canada (EIC) immigration file. f_c factors from the EIC immigration file are intermediate between those derived from interprovincial immigrant data and U.S. emigrant data. The figure based on the f_c factor from the immigration file (275,762) is higher than the official estimate of emigrants (235,481), but is close to that derived from the 1986 Reverse Record Check study (288,376). If the official estimate of the number of emigrants were increased to 275,762, the 1986 error of closure between the population estimate and census counts would be reduced from 0.95% to 0.79%.

In sum, for the 1981-86 period the estimates of emigrants seemed to be improved by taking f_c factors from the Canada Employment and Immigration (EIC) immigrant file rather than the Revenue Canada tax file.

Yet in March 1989, it was discovered that emigrant estimates based on Family Allowance files and an f_c factor derived from EIC immigration data were still too low after 1986. This seems to be a result of the high proportion (33%) of Canadian emigrants destined for the U.S. from 1981 to 1986, according to U.S. data.

An analysis was also made of a method combining U.S. Department of Justice, Immigration and Naturalization Service data on the numbers emigrating to the U.S. from Canada; child emigrant counts (ages 0-17) from Family Allowance files and an f_c factor obtained from the EIC immigration file for all countries other than the U.S. For 1981 to 1986, the estimated number of emigrants by this method was 285,413. This revised estimate is much closer to the Reverse Record Check study figure (288,376).

Table 3

Estimates of Net Interprovincial Migration from 1986 Census Data on Mobility, Family Allowance Files, Income Tax Files, and Residual Method, Canada, Provinces and Territories, 1981-1986

Geographic Area	1986 Census[1]	Family Allowance Files	Income Tax Files	Residual Method[2]
CANADA	0	0	0	−238,178
Nfld.	−16,550	−14,837	−15,051	−26,111
P.E.I.	1,540	293	751	−509
N.S.	6,275	5,204	6,895	−4,095
N.B.	−1,370	−2,239	−65	−11,212
Que.	−63,295	−76,040	−81,254	−167,286
Ont.	99,355	115,497	121,767	57,147
Man.	−1,555	−3,700	−2,634	−8,180
Sask.	−2,820	−668	−2,974	−13,564
Alta.	−27,665	−34,073	−31,676	−50,811
B.C.	9,500	13,289	7,382	−12,418
Yukon	−2,665	−2,381	−2,775	−1,643
N.W.T.	−755	−345	−366	504

[1] Population 5 years of age and over.
[2] The residual method for estimating net interprovincial migration is:

$$\text{Net Migration} = \text{Growth of Census Population between time } t \text{ and } t + 5 - [(\text{Births} - \text{Deaths}) + (\text{Immigration} - \text{Emigration})].$$

Source: Demography Division, Statistics Canada.

3.2 Interprovincial Migration Data

To test the accuracy of estimates of interprovincial migration obtained from the Revenue Canada tax file, two evaluations were conducted: (i) a comparison of sets of interprovincial migration data derived from the Revenue Canada tax files and Family Allowance files; and (ii) a comparison of the errors of closure of population estimates for two sets of internal migration data.

Table 3 presents net interprovincial migration estimates derived from four sources: 1986 Census data on mobility; the Revenue Canada tax file; the Family Allowance file; and the residual-based net migration estimate. For all provinces, estimates of internal migration derived from the 1986 Census mobility data, the Revenue Canada tax file and Family Allowance files were consistent on the direction of net migration. All sources except the residual-based method show positive net migration for Prince Edward Island, Nova Scotia, Ontario and British Columbia. In other provinces, net migration was negative.

The estimates of net interprovincial migration from Family Allowance files and Revenue Canada tax files are not strictly comparable to the residual method. By definition, the sum of net interprovincial migration in Canada, should be zero. However, the sum produced using the residual method is about 238,000. In addition, the differences between the residual-based and the Revenue Canada/Family Allowance-based net interprovincial migration estimates are very high in Newfoundland, New Brunswick, Quebec, Ontario and Alberta.

The coefficient of variation (the ratio of the standard deviation of the average absolute error of closure for the provinces to the average absolute error of closure) was used to measure the relative accuracy of the internal migration estimates. The other estimates of the components of population change were assumed to be accurate. Statistically, a coefficient of variation of 20% to 30% is normally acceptable.

Table 4

Error of Closure Between Alternative Population Estimates and Census Counts by Province and Territory 1971, 1976, 1981 and 1986

Geographic Area	Error of Closure[1] (%)							
	1971		1976		1981		1986	
	Income Tax	FA	Income Tax	FA	Income Tax	FA	Income Tax	FA
Newfoundland	−2.08	−1.64	0.49	1.34	1.63	2.30	1.97	2.01
Prince Edward Island	−2.09	−2.01	0.17	2.11	−0.05	1.02	0.99	0.63
Nova Scotia	−1.68	−2.39	−0.20	1.18	0.30	0.40	1.24	1.04
New Brunswick	−1.93	−2.65	−1.29	1.81	0.13	0.54	1.58	1.04
Quebec	−0.33	−0.97	−0.05	−0.18	−0.30	−0.07	1.32	1.40
Ontario	0.11	0.99	0.15	0.16	0.64	0.37	0.72	0.65
Manitoba	0.29	0.38	−0.27	0.39	1.07	0.87	0.51	0.41
Saskatchewan	0.44	−0.33	0.45	0.37	−0.31	0.28	1.08	1.31
Alberta	−0.14	0.52	−1.07	−1.11	−2.39	−2.64	0.73	0.63
British Columbia	0.01	−1.34	0.28	−1.10	0.03	−0.07	0.59	0.79
Yukon	−5.36	−5.99	−0.87	3.79	−1.98	2.06	−4.78	−3.10
Northwest Territories	−2.12	2.64	−12.98	−3.39	−7.08	0.43	−1.44	−1.40
Average Absolute Error								
10 provinces	0.91	1.33	0.44	0.97	0.69	0.86	1.07	1.01
Provinces and Territories	1.38	1.82	1.52	1.41	1.33	0.92	1.41	1.22

Note: 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+.

[1] Error of closure is calculated using the following equation:

$$\text{Error of closure} = \left(\frac{\text{Estimate} - \text{Census count}}{\text{Census count}}\right) \times 100$$

Income Tax: Revenue Canada Income Tax File. FA: Family Allowance File.

Source: Estimates of interprovincial migration based on Family Allowance data, Demography Division, Statistics Canada.
Estimates of interprovincial migration based on tax data, Small Area and Administrative Development Division, Statistics Canada.

Table 5

Coefficients of Variation of the Average Absolute Error of Closure between the Population Estimates and Census Counts among Provinces (n = 10), by Source of Interprovincial Migration Estimates, 1966-1971, 1971-1976, 1976-1981 and 1981-1986

Period ($t,t+5$)	Source	AAE ($t+5$)	Standard Deviation	Coefficient of Variation (%)
		(1)	(2)	(3) = (2 ÷ 1) × 100
1966-1971	Income Tax	0.91	0.2863	31
	FA	1.33	0.2642	20
1971-1976	Income Tax	0.44	0.1317	30
	FA	0.97	0.2135	22
1976-1981	Income Tax	0.69	0.2463	36
	FA	0.86	0.2855	33
1981-1986	Income Tax	1.07	0.1496	14
	FA	1.01	0.1570	16

Note: AAE: Average absolute error of closure.
Income Tax: Revenue Canada Income Tax File.
FA: Family Allowance File.

Source: Demography Division, Statistics Canada.

However, one could argue that the coefficient of variation is not a good indicator of the quality of internal migration data. For example, a set of estimates with an absolute error of closure of 10% for every province would give a coefficient of variation of zeros and consequently would be preferable to a set of estimates with closure errors ranging between -1.0% and 1.0%. For cases like this, a quality measure that takes into account the size of the absolute error of closure as well as the standard deviation of absolute closure errors is clearly required. However, the likelihood of the provinces having the same absolute error of closure is extremely low (see Table 5), hence, the application of the coefficient of variation in this paper seemed to be valid.

Table 5 shows the coefficient of variation (computed from figures in Table 4) for population estimates based on two sets of internal migration estimates and the census counts for 1971, 1976, 1981 and 1986. Before 1976, the coefficients of variation for migration data from tax files were 50% higher for data from the Family Allowance file. This was expected, since the method for estimating migration from tax files was in the developmental stage. Futhermore, in estimating the number of interprovincial migrants, the f_j factor (adult to child migration rates) was based on Census mobility data, an approach found to be less satisfactory than the current method. However, for 1976-1981 and 1981-1986, the gap in the coefficient of variation between the tax and Family Allowance migration data narrowed considerably.

The tax-based migration data coefficient of variation was 9% higher in 1981 and 12% lower in 1986 than the coefficient of variation based on the Family Allowance file. Hence, the two sets of data are comparable, producing similar provincial estimates and errors of closure with the same level of variation among provinces. Since the coefficient of variation for each set is under 20%, they provide acceptable data on internal migration.

In conclusion, estimates of interprovincial migration from the Revenue Canada tax files for 1981-1986 are consistent with estimates from the Family Allowance file. By province, they yield small variations in the errors of closure.

4. CONCLUSION AND DISCUSSION

The Family Allowance files and Revenue Canada tax files play important roles in providing consistent emigration and internal migration estimates for Canada, and for the provinces and territories. For 1981 to 1986, estimates of emigrants and interprovincial migrants obtained from these files are acceptable for estimating total population.

Nationally the error of closure (the difference between the population estimates and census counts) for 1986 was higher than for the census years 1971, 1976 and 1981. In addition, the errors of closure by province in 1986 were positively biased, indicating that in all provinces the estimates were higher than census counts.

These discrepancies are largely a result of differences in coverage of the 1981 Census population, which was used as the bench-mark, and coverage of the 1986 Census population. The Reverse Record Check estimate of the 1981 undercoverage rate for Canada was 2.01%. The estimate for the 1986 Census was considerably higher, 3.21%.

Errors in the estimates of the other components of change may also partly account for the discrepancies.

ACKNOWLEDGEMENT

We are grateful to the reviewers of this article for their comments and suggestions.

REFERENCES

BEAUJOT, R., and RAPPAK, J.P. (1988). *Emigration from Canada: its Importance and Interpretation.* Ottawa: Employment and Immigration Canada.

NORRIS, D., BRITTON, M., and VERMA, R.B.P. (1982). The use of administrative records for estimating migration and population. *Statistics of Income and Related Administrative Record Research: 1982*, Washington, D.C.: Department of the Treasury, Internal Revenue Service.

NORRIS, D., and STANDISH, L.D. (1983). A technical report on the development of migration data from taxation records. Technical Report, Small Area and Administrative Data Development Division, Statistics Canada.

RABY, R., MARTEL, J., and CARTIER, G. (1989). Issues in the current postcensal population estimates. Paper presented at the Federal-Provincial Committee on Demography, Ottawa.

STATISTICS CANADA (1987). *Population Estimation Methods, Canada.* Catalogue 91-528E, Statistics Canada.

VERMA, R.B.P., and PARENT, P. (1985). An overview of the strengths and weaknesses of the selected administrative data files. *Survey Methodology*, 11, 171-179.