THE REGRESSION ESTIMATES OF POPULATION FOR SUB-PROVINCIAL AREAS IN CANADA[1]

Ravi B.P. Verma, K.G. Basavarajappa,
and
Rosemary K. Bender[2]

In order to improve the timeliness, accuracy and consistency of population estimates for different geographic areas, Statistics Canada has developed new methods of estimation for sub-provincial areas (census divisions and census metropolitan areas). Beginning with 1982, two sets of population estimates (regression and component based) will be published yearly, appearing 3-4 months and 12-15 months, respectively, from the reference date.

The regression technique uses family allowance recipients as the main symptomatic indicator and where available, additional indicators - reference population from provincial health insurance files and hydro accounts - to derive population change for the current year. The first set is obtained by adding this change to the second set for the previous year produced by the component method, with births and deaths from vital registers, and estimated migration from Revenue Canada taxation files. The two sets were found to be statistically similar with respect to accuracy, though the first set is more timely, and the second provides more details on the components of population change.

1. INTRODUCTION

Annual estimates of population for sub-provincial areas such as Census Divisions and Census Metropolitan Areas are useful for the planning of housing, hospitals, schools, colleges and universities and other social service programmes, studies of labour market areas, allocation of funds, computing vital

---

[1] The earlier version of this paper was presented at the Joint Statistical Meetings of the American Statistical Association, the Biometric Society, the Institute of Mathematical Statistics and the Statistical Society of Canada in Toronto, August 1983.

[2] Ravi B.P. Verma, K.G. Basavarajappa and Rosemary K. Bender, Demography Division, Statistics Canada.

rates, designing some surveys, computing the index of economic disparities, forecasting the number of tax payers etc. In particular, these estimates are required for weighting the results of Labour Force and Consumer Income and Expenditure Surveys and by the Department of Regional Economic Expansion, Revenue Canada and some provincial governments.

In order to improve the timeliness, accuracy, and to maintain the consistency among population estimates for different geographic areas, Statistics Canada has recently devised new methods for estimating the total population for census divisions and census metropolitan a eas. The objectives of this paper are to describe the post 1981 methodology for estimating the total population for these sub-provincial areas, discuss the accuracy of the methodology, review the work done by the provincial governments, and to discuss some factors which have played a vital role in the selection of some of these methods.

2. POPULATION ESTIMATION PROGRAMME FOR THE POST 1981 PERIOD

Beginning with 1982, for each year, Statistics Canada will publish postcensal population estimates for census divisions at two different periods. This is similar to the established practice for census metropolitan areas.

The first set which is based on a combination of regression model and component procedure and which provides no details on components of population change, will be published near the end of September of each year, i.e., 4 months from the reference date. Those estimates are termed regression-nested estimates (see Table 1).

As can be seen from Table 1, the first set of population estimates as of June 1, 1982 are the regression estimates, but for other years 1983 to 1986 they will be obtained by adding the change in the regression estimates to the second set of population estimates (obtained by the component procedure) for the preceding year. This approach ties together the two sets of the postcensal estimates for a specific reference date.

Specifications of the regression method by province, for census divisions and census metropolitan areas are presented in Table 2. For Census Divisions in each province, regressions will be based on the best available symptomatic indicators of population change namely, the number of family allowance beneficiaries aged 1-14, reference population taken from health insurance files (Saskatchewan and Alberta), and the number of hydro accounts (British Columbia). Regression models are primarily designed to maximize the accuracy of population estimates. For census metropolitan areas, the first set of population estimates for the post 1981 period will take input from the regression (Ratio-correlation) method using family allowance recipients aged 1-14 as the symptomatic indicator of population change. The form of regression as well as the variables selected are based on the criterion of minimum average absolute error (defined in Table 2) for alternate estimation methods for the 1976-81 period. These new methods are expected to result in estimates that are more reliable than those actually produced by alternate procedures for the intercensal years between 1976 and 1981.

The second set of estimates, produced using the component method, will provide details on each of the components of population change, and would be published about 12-15 months following the reference date. Birth and death data will be obtained from the vital statistics records, and the migration data will be derived from Revenue Canada Tax Files (Norris and Standish, 1983). The components of international migration derived from Revenue Canada Files, however, need further adjustment. For census divisions, this adjustment will be based on the immigration data eminating from the department of Employment and Immigration, and the independent estimate of emigration derived by Statistics Canada using the Family Allowance Files (Raby and Parent, 1982). For census metropolitan areas, no adjustment is needed for the immigration data, as they will be compiled from the intended destination of immigrants to CMA, from the department of Employment and Immigration. But the adjustment is applied to the estimates of emigrants which are derived as in the case of CDS.

For the first set of postcensal population estimates for census divisions by province, a further adjustment (prorating) is made to make them consistent with the corresponding provincial population totals. This is not necessary

for the second set, as each of the components of population change is already consistent with the corresponding provincial total. Similarly, an adjustment is made only to the first set of postcensal population estimates for census metropolitan areas.

3. EVALUATION OF ESTIMATION METHODS

The relative accuracies of the regression method, the methods used during the period 1976-81, and the component method using migration data from Revenue Canada Tax Files are examined elsewhere (Norris, Britton and Verma, 1982). The accuracy is measured by comparing the 1981 estimates constructed from the 1976 base population with the enumerated totals as obtained from the census of 1981.

Methods of estimation are evaluated with respect to three criteria: (i) greater accuracy, (ii) timeliness and (iii) consistency. As mentioned above, accuracy of population estimates is measured by comparing the estimates with the census counts by computing the indices of average absolute error and index of misallocation. The error is defined as the difference between the estimate and the census count. Average absolute error is the arithmetic mean of percentage errors disregarding sign (see Table 2). Index of misallocation is the index of dissimilarity, defined as half of the sum of absolute differences of the two sets of percent distribution of population estimates. Timeliness refers to the availability of estimates within as short a time as possible after the reference date. Consistency refers both to data consistency in the estimation of population being developed at various area levels of disaggregation using the same data source, and to uniformity in the methods of estimation. It must, however, be recognized that in some cases, the use of different data sources and different methods may be unavoidable.

3.1 Census Divisions

Relative accuracies of four sets of estimates as of June 1, 1981 obtained by

four different methods for census divisions are presented in Table 3. It appears that each of the alternative methods (regression, regression-nested, and component) is superior to the old methods used during the period 1976-81. For Canada as a whole, among the alternative methods, regression-nested, derived by adding the change between two year regression estimates to the previous year component estimate, seems to be the most accurate with a lowest average absolute error of 1.7%. Betweeen regression and component methods, the regression-direct is observed to be less accurate. This is true in all provinces except the province of Saskatchewan in which the regression estimates are based on the reference population from health insurance files as the indicator of the population change. The accuracy of this indicator in estimating population is very high as indicated by a very low average absolute error, 1.43 percent. In 5 out of 10 provinces the regression-nested is slightly more accurate than the component method.

In order to assess the relative accuracy of each of the alternative methods, the paired t-test was also carried out. For Canada as a whole, it was found that the differences were statistically significant between the estimates obtained from the regression-direct and component method. This is also true in Ontario and Quebec. In contrast, the differences were not statistically significant between the regression-nested and the component method indicating that both these methods are statistically similar in terms of the accuracy. Similar results were observed when the t-test was performed on the weighted average of absolute errors which takes into account size of population.

3.1.1 Temporal Stability

In order to illustrate the temporal stability of three sets of postcensal estimates for census divisions (regression, component and regression-nested), the index of dissimilarity was computed for the years 1977 to 1981 and is presented in Table 4. It may be observed that both the disparities between the regression and component estimates (A) and the regression and nested estimates (C) increase over time. However, the disparity between the regression-nested and component estimates fluctuates slightly and is found to be minimum. Thus these two methods, in general, provide similar results

during 1976 to 1981.

The component and regression methods are independent and so the results may be expected to diverge, whereas, the regression-nested and the component methods overlap and so the results tend to be similar.

The largest gap between the regression and the component estimates is not expected to fall, because there are some theoretical weaknesses inherent in the regression method. For example, the model may fit well for the previous time period, but may predict poorly during the succeeding period. The assumption in the regression method that the vector of regression coefficients for symptomatic indicators is invariant from the immediately preceding intercensal period to the postcensal period is often questionable. In practice, this invariance may not hold good over time, both because of structural changes in the underlying relationships of the variables, and also because of the improvement in the quality of the symptomatic indicators over time.

3.1.2 The Effects of Structural Changes

In order to examine the effects of structural changes on the differences between the 1976 and 1981 average errors, the 1981 average errors resulting from the equations of the model 1971-76 were compared with those resulting from the regression equations of the model period, 1976-81. It may be seen from Table 5 that the 1981 average errors resulting from the equations for two different time periods are quite comparable in all provinces except Saskatchewan, where the error declined by nearly 50% from 1.3% to 0.7%.

Due to structural changes, the relationship between the variations in symptomatic indicators (vital events and family allowance) and variations in population have undergone changes during the period, 1976-81. This is probably so for the provinces Quebec, Manitoba and Alberta. During the period 1976-81, the characteristics of the people moving from the eastern and maritime provinces to the western provinces may have changed considerably. For example, as the family allowances are limited to the families with

children, movement of single persons and families without children were not captured by the changes in the family allowance indicator. Due to this, the family allowance recipients as an important predictor of the population change in the regression model, 1976-81 failed to predict adequately. Thus, it is clear that the average errors in 1981 resulting from models of both time periods, 1971-76 and 1976-81 were high, because of structural changes.

A part of the difference in the average errors between 1976 and 1981 is also due to changes in the quality of family allowance data. The numbers of family allowance recipients are produced at the census division level by converting postal codes to standard geographic codes. In 1976, the conversion file had problems of missing and overlapping postal codes. In particular the percentage of missing codes in 1976 was high in maritime provinces and Ontario. But, by 1981, the magnitude of the problem of missing postal codes in the FA files had declined in all provinces. Thus, the change in the quality of the family allowance indicator between the years 1976 and 1981 may have also affected the quality of the regression coefficients during the period 1976-81.

3.2 CMAs and the Non-CMA Unit

Table 6 presents the average absolute errors for CMAs according to three types of estimates, viz., regression-direct, nested and component. It may be seen that the component method provides estimates with the lowest errors at Canada level. The regression nested procedure comes second best. The same findings as for census divisions hold good when we consider the indices of dissimilarity which are given below:

Nested vs. Component	0.98%
Regression-direct vs. Component	1.15%
Regression-direct vs. Nested	1.09%

3.2.1 Consistency and Timeliness

In terms of the accuracy of population estimates and consistency with respect to sources of input data and methods used for estimating the population of

different geographic areas, (provinces and territories, CDs and CMAs), the component method appears to be the most suitable. In addition, the component method provides more detailed and consistent information on components of population change, e.g., consistent set of internal migration figures classified by streams (in- and out-) and by broad age groups and sex for the province and its sub-provincial areas. However, this method does not provide timely estimates. The delay is expected to be about 12-15 months. The proposed regression method using family allowance recipients and/or other symptomatic indicators on the other hand, can provide estimates with a delay of about 3-4 months.

From Table 6, it may also be seen that in terms of the accuracy of the population estimates, the component and the regression-nested are closer to each other than regression-direct and component. But in terms of the timeliness, the regression-nested is superior to the component method.

4. ALTERNATIVE METHOD TESTED

The type of regression method by province shown in Table 2 is the most accurate for a given province among several alternative methods of estimation which were tested over the period 1976-81. These methods are: two types of component methods using migration estimates from school enrolment data and tax files, vital rates method, ratio method using the provincial administative files, proportional allocation method based on family allowance recipients, and six types of regression methods (ratio-correlation, weighted ratio-correlation, ridge weighted ratio-correlation, difference-correlation, weighted difference-correlation, and ridge weighted difference-correlation). Of these methods, the methods used for official estimates during 1976-81 include the component method using migration estimates from school enrolment, ratio method and ratio-correlation method (Dominion Bureau of Statistics, 1967).

Weighted regression method was adopted in order to control for heteroscedasticity. In this procedure, we transform the data set with the calculated weights such that one obtains a random error term (e) with constant variance.

We have used the Goldfield-Quandt procedure for testing the assumption of homoscedasticity (Johnston, 1963). Ridge regression controls for multi-collinearity. 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).

The accuracy of all these methods of estimation are thoroughly evaluated and the results presented in the three reports by Verma, Basavarajappa and Bender (1982a, 1982b, 1982c).

5. BACKGROUND HISTORY

In adopting the post 1981 methodology for estimating the population for sub-provincial areas, the following points were considered: the accuracy of the methods used during the period 1976-81, theoretical issues in the regression method, review of the work done by the provincial governments, two sets of official estimates for certain census divisions - one produced by Statistics Canada and the other produced by some provincial governments, consideration for a small area data development project and demand by other private users. A brief discussion of some of these points is given in the following paragraphs.

5.1 Review of Methodology Used during 1976-81 for Census Divisions and Census Metropolitan Areas.

Methods used during 1976-81 census divisions were specific to the provinces as presented in Table 3. These methods had many limitations (Verma and Basavarajappa, 1982a). These included the inadequacies of symptomatic indicators in capturing the current population changes (e.g., births and deaths), excessive time lag of about 2 years (due to delay in obtaining data on school enrolments) and some specification problems. The latter arose because in some provinces, particularly in those with large rural areas, school enrolment may not conform precisely to residential patterns due to transportation of children across census division boundaries. As a result of these limitations, the accuracy of the estimates for census divisions became

unsatisfactory.

During the period 1976-1981, the component method was used to produce estimates for census metropolitan areas in Canada using births and deaths from Vital Statistitcs registers and Immigrants to CMAs from Employment and Immigration Department. The accuracy of the population estimates for CMAs was unsatisfactory primarily due to the weaknesses in the methodology for estimating emigration and internal migration for which no direct sources were available (Catalogue No. 91-207).

5.1.1 Ratio vs. Difference Correlation Methods.

Schmitt and Crosetti and many others have claimed that the ratio-correlation method is one of the most accurate methods (using as the criterion the Average Absolute Error - AAE) (Balakrishnan, 1960; Goldberg, Rao and Namboodiri, 1964; Swanson, 1978; N.R.C., 1980; Mandell and Tayman, 1982). Later, some researchers including Schmitt and Grier suggested that the difference-correlation method is an improvement over the ratio-correlation method (Schmitt and Grier, 1966; O'Hare, 1976). This was because the difference-correlation method produced constant mean, a lower mean square error (M.S.E.), higher intercorrelation among the variables, and a resulting higher squared value of the coefficient of multiple correlation (R^2). These features are often used to evaluate the fitting of a regression model and are considered desirable.

However, no consistent relationship between the higher R^2 and the average absolute error has been observed. The accuracy of population estimates produced by the regression method is highly dependent on the temporal stability of the regression coefficients. In this respect, a recent study has shown that the ratio-correlation method was more suitable than the difference-correlation method (Mandall and Tayman, 1982). The difference-correlation method produced a higher multi-collinearity than the ratio-correlation. Due to this, the difference-correlation shows higher instabibility in the regression coefficients over time-periods (Spar and Martin, 1979).

A review of both techniques has revealed that neither the ratio-correlation,

nor the difference-correlation method uniformly or routinely outperforms the other (O'Hare, 1980). This was also confirmed by Verma, Basavarajappa and Bender (1982a).

In light of the above findings, a multiple-model frame work seems to be the most appropriate course for evaluating the competing estimation techniques. In fact, this is what has been employed in the present estimation programme.

5.1.2 Review of the Work done by Provincial Governments

A survey of provincial/territorial agencies producing population estimates and projections revealed that neither the methods nor the geographic divisions for which the estimates were produced were uniform. Some prepared estimates for census divisions and other areal units, non prepared estimates for census metropolitan areas.

To estimate populations of census divisions, or counties, one popular approach adopted by Ontario, Alberta and Northwest Territories is the component method described earlier.

The Northwest Territories obtain births and deaths from its Bureau of Vital Statistics. It estimates net migration with a time related cohort model for the population subgroup 1-14 years of age using family allowance recipients and school enrolment data.

Ontario also uses birth and death data from their Bureau of Vital Statistics. However, it estimates net migration from the changes of addresses from the drivers licence files of the Ministry of Transport.

Alberta uses a combination of two techniques. A ratio-correlation method estimates population change using births, school enrolment and the provincial health insurance plan as symptomatic indicators. Using the component method approach, the net migration is then obtained as a residual of the regression-based population growth, and births and deaths.

British Columbia, on the other hand, uses a combination of the difference-correlation method, with hydro billings, family allowance and vital statistics as symptomatic indicators, and the proportional allocation method. It is the only province that adjusts its subprovincial estimates so as to correspond to the provincial total published by Statistics Canada.

Quebec uses the best of several techniques to estimate their municipal populations. One method uses the figures provided by the municipalities and if found reasonable, these have priority over all others. Other methods use the rates of growth in hydro billings in combination with estimates/counts of the preceding year.

Newfoundland also estimates their communities using hydro billings as a source of input data. It combines this with preceding census counts, number of households, and average number of people per household.

Manitoba estimates the population of its municipalities in much the same way as Saskatchewan does at the provincial level. The count of elegible persons registered under its provincial medical health insurance plan, along with appropriate adjustment factors, is used to directly estimate the municipal populations in these two provinces.

However, no systematic evaluation of these estimates is available. Newfoundland, Quebec and Ontario are in the process of evaluating their estimates, British Columbia's evaluation of their estimates supports the continuation of their estimation methodologies for the post 1981 period.

The time lag after the reference date for which estimates become available ranges up to six months. Manitoba and Saskatchewan produce data within two months, the Northwest Territories, British Columbia and Ontario within four months and Quebec and Newfoundland within six months.

In conclusion, there is no uniformity of methods across the country. Each provincial/territory uses techniques that suit its particular needs, and which take advantage of provincial administrative data files.

5.1.3 Federal-Provincial Consultations

The new techniques devised for estimating the population of sub-provincial areas were discussed at meetings of the Federal-Provincial Committee on Demography. It is well to remember that the regression method was devised primarily for providing timely preliminary totals and the updating of these is firmly anchored in the component method. The question of the usefulnes of figures for CDs and CMAs is also worth considering. While the provinces need population estimates for municipal and administrative regions more than for CDs and CMAs, the latter are needed for Statistics Canada's internal uses and as building blocks for specified areas. Over the years, it has been observed that there has been a sufficient demand for estimates for CDs and CMAs. The lack of resources is also an important factor in preventing the extension of the estimation procedures for small sub-divisions of the provinces. Because of this, with some technical assistance from Statistics Canada, some provinces are planning to undertake the task of preparing population estimates at the municipal and other smaller divisions.

It may be noted that the above arrangement also avoids the duplication of efforts by the provincial and federal governments relating to the preparation of estimates for provinces and sub-provincial areas.

6. EVALUATIVE DISCUSSION

The research during the past year, carried out in collaboration with several provincial statisticians, resulted in the development of improved methods for estimating the population of census divisions and census metropolitan areas. As of 1982, for each year, Statistics Canada will publish two sets of post-censal population estimates for sub-provincial areas at two different periods.

The first set which is based on a regression model (and which refers to June 1 of each year) will be published near the end of September of each year, i.e., with a delay of utmost 4 months. The second set of estimates referring to the same date, produced by the component method using migration data derived from Revenue Canada Taxation Files, and the numbers of births and deaths from Vital

Registers, will be published about 12-15 months following the reference date.

These new methods are expected to result in estimates that are more reliable than those actually produced for the intercensal years between 1976 and 1981. These, more accurate and timely sub-provincial population estimates will be crucial to the Small Area Data Development Program that has just been launched by Statistics Canada.

It should be realized that the types of regression method that gave rise to a satisfactory pattern of error during 1976-81 for each province may turn out to be unsatisfactory during 1981-86, thereby giving rise to estimates with higher errors than anticipated. For example, on average, the regression model error was 2% in 1976 but when coefficients of the 1971-76 were applied to produce the estimates in 1981, the accuracy of the 1981 population estimates as compared to the 1981 census counts for census divisions was found to be 2.54%. Thus, we anticipate that the error as shown in Table 2 may increase by about 0.50 percentage points. However, the error in 1986 for the regression-nested estimate derived by adding the change in the regression estimates to that obtained by component method is expected to be very close to that of the component method.

One might argue that the practice of changing one set of estimates with another set of estimates for a specific reference date will have a negative impact on the planning for different social programmes. Also, if the two sets do not differ from each other very much, is there any need for producing both sets? The defense is that the first provided timely data and of acceptable quality, and the second, besides providing the relatively more detailed information on the components, provides estimates of acceptable and perhaps better accuracy.

ACKNOWLEDGEMENTS

We are thankful to Mr. Gordon Brackstone, Mr. Garnet Picot and Mr. Doug Drew of Statistics Canada, Dr. N.M. Lalu of University of Alberta, and the referee

as well as the assistant editor of this journal for many helpful comments.

TABLE 1

Methodology for the First Set of Population Estimates (Regression-Nested) for Census Divisions and Census Metropolitan Areas

Time	Regression Estimate	Component Estimate*	Regression-nested Estimate
t	P_t	P_t^c (census)	P_t
t+1	P_{t+1}	P'_{t+1}	P_{t+1}
t+2	P_{t+2}	P'_{t+2}	$P'_{t+1} + [P_{t+2} - P_{t+1}]$
t+3	P_{t+3}	P'_{t+3}	$P'_{t+2} + [P_{t+3} - P_{t+2}]$
t+4	P_{t+4}	P'_{t+4}	$P'_{t+3} + [P_{t+4} - P_{t+3}]$
t+5	P_{t+5}	P'_{t+5}	$P'_{t+4} + [P_{t+5} - P_{t+4}]$

* The method uses births and deaths from Vital Registration Records and migration data from Revenue Canada Taxation Files.

TABLE 2

Specifications of the Regression Method by Province for Estimating the Population Totals for Census Divisions and Census Metropolitan Areas, Post 1981 Period

Area/Province	Type*	Model Period	Symptomatic Indicator	Test 1981 AAE
Census Divisions				
Nfld. - P.E.I.	RC	1976-81	F	1.27
N.S.	RC	1971-76, 1976-81	F	1.50
N.B.	RC	1976-81	F	1.30
Quebec	RC	1976-81	F	1.81
Ontario	RC	1976-81	F	1.99
Manitoba	WDC	1971-76, 1976-81	F	3.13
Saskatchewan	DC	1976-81	CP	0.62
Alberta	WRC	1976-81	F, HC	1.89
B.C.	WDC	1971-76, 1976-81	F, Hydro	2.14
TOTAL				1.84
CMAs	RC	1976-81	F	2.30

Note: F: Family Allowance Recipients aged 1-14 years old.
CP: Covered Population.
HC: Health Care Files.

AAE: Average Absolute Error $= \frac{1}{N} \Sigma \left| \frac{E_i - P_i}{P_i} \right|$.

E_i: Estimated Population for Census Divisions.

P_i: Census Population for Census Divisions.

N: Number of Census Divisions with Province.
RC: Ratio-correlation.
WDC: Weighted-Difference correlation.
WRC: Weighted-Ratio-correlation.
DC: Difference-correlation.
CMAs: Census Metropolitan Areas.

* For a description of the types of regression methods, the readers are referred to the paper by W. O'Hare [10].

TABLE 3

Evaluation of Population Estimates, June 1, 1981
(Average Absolute Error)

Province	No. CDs.	Regression Direct (1)	Regression Nested	Component	Old Method Used (2)
NFLD. - P.E.I.	13	1.36	0.67	1.00	2.6
N.S.	18	1.64	1.27	1.07	6.8
N.B.	15	1.59	1.05	1.06	3.3
Quebec	76	3.10	1.63	2.02	2.5
Ontario	53	2.17	1.26	1.21	1.5
Manitoba	23	3.33	2.57	2.58	4.4
Saskatchewan	18	1.43	1.96	2.10	2.0
Alberta	15	4.45	2.84	2.39	5.1
B.C.	29	2.45	2.50	2.39	9.2
TOTAL	260	2.55	1.72	1.80	2.9

Notes: (1) The method uses as symptomatic variables reference population for Saskatchewan and family allowance recipients for other provinces.

The model period for all provinces in 1971-1976, using weighted ratio correlation for Alberta, weighted difference correlations for British Columbia, and ratio correlation for all other provinces.

(2) Methods used during 1976-81: Component II: Prince Edward Island, Nova Scotia, New Brunswick, Manitoba, Alberta and British Columbia.

Ratio Method: Ontario and Saskatchewan.

Ratio-correlation: Newfoundland and Quebec.

For a description of all these old methods, the readers are referred to the Statistics Canada Catalogue No. 91-206 [15].

TABLE 4

Temporal Stability of Three Sets of Postcensal Estimates for Census Divisions (Regression-direct, Regression-nested, Component) 1977-1981

Provinces		1977	1978	1979	1980	1981
NFLD.	A.	0.17	0.33	0.41	0.34	0.51
	B.	0.17	0.19	0.19	0.13	0.13
	C.	0.00	0.17	0.34	0.35	0.41
P.E.I.	A.	0.17	0.26	0.25	0.51	0.51
	B.	0.17	0.08	0.19	0.02	0.24
	C.	0.00	0.17	0.26	0.52	0.26
N.S.	A.	0.29	0.53	0.60	0.63	0.64
	B.	0.29	0.30	0.18	0.23	0.19
	C.	0.00	0.53	0.38	0.45	0.70
N.B.	A.	0.52	0.38	0.46	0.71	0.48
	B.	0.52	0.48	0.44	0.52	0.37
	C.	0.00	0.53	0.38	0.45	0.70
QUE.	A.	1.02	0.64	0.81	0.99	1.13
	B.	1.02	0.72	0.27	0.57	0.54
	C.	0.00	1.05	0.66	0.80	0.98
ONT.	A.	1.69	0.58	0.70	0.99	0.94
	B.	1.69	1.75	0.31	0.49	0.56
	C.	0.00	1.67	0.55	0.71	0.96
MAN.	A.	0.21	0.39	0.60	0.70	0.80
	B.	0.21	0.26	0.26	0.21	0.19
	C.	0.00	0.20	0.42	0.59	0.70
SASK.	A.	0.37	0.52	0.53	0.70	0.78
	B.	0.37	0.18	0.26	0.25	0.18
	C.	0.00	0.38	0.51	0.55	0.68
ALTA.	A.	0.45	0.45	0.57	0.89	1.18
	B.	0.45	0.21	0.27	0.41	0.36
	C.	0.00	0.44	0.43	0.56	0.86
B.C.	A.	0.39	0.45	0.76	0.95	0.93
	B.	0.39	0.32	0.41	0.23	0.29
	C.	0.00	0.37	0.43	0.76	0.94

Note: Index of dissimilarity between estimates E_1 and E_2 for a province with n census divisions and total population P is given by:

$$\frac{1}{2}\sum_{i=1}^{n}\left|\frac{E_{1i} - E_{2i}}{P}\right|$$

A: Index of dissimilarity between regression-direct and component estimates.
B: Index of dissimilarity between regression-nested and component estimates.
C: Index of dissimilarity between regression and regression-nested estimates.

Source: Demography Division, Statistics Canada, February 1983.

TABLE 5

Comparaison of the Accuracy of the Regression Methods for the Model Periods 1971-76 and 1976-81

	Regression		Model 1971-1976		Model 1976-81
	Type	Indicator	Test 1976 AAE	Test 1981 AAE	Test 1981 AAE
Nfld. - P.E.I.	RC	F	1.6	1.4	1.3
N.S.	RC	F	1.8	2.0	1.6
N.B.	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	CP	1.5	1.3	0.7
Alberta	RC	F	3.1	4.6	4.2
B.C.	WDC	F	3.1	4.0	2.3
CANADA			1.96	2.54	2.04

Note: WDC: Weighted difference correlation.
RC: Ratio correlation with ordinary least square.
F: Family allowance recipients.
V: Vital events (Births + deaths).
CP: Covered population in Saskatchewan.
AAE: Average absolute error.

TABLE 6

Evaluation of 1981 Population Estimates
(CMAs and Non-CMA)

Method	Average Absolute Error (%)
Regression (F), (1971-76)	2.25
Regression-Nested	2.21
Component (Tax)	1.47

Note: F: Family Allowance Recipients Aged 1-14 years.

$$\text{Average Absolute Error} = \frac{1}{n} \Sigma \left| \frac{\text{Estimate - Census}}{\text{Census}} \right| \times 100.$$

n = Number of CMAs and non-CMAs.

REFERENCES

[1] Dominion Bureau of Statistics (1967). Population Estimates for Counties and Census Divisions, Catalogue No. 91-206, Ottawa: The Queen's Printer.

[2] Goldberg, D. and Balakrishnan, T.R. (1960). A Partial Evaluation of Four Estimation Techniques. Paper presented at the Annual Meeting of the Social Statistics Section, American Statistical Association.

[3] Goldberg, D., Rao, V.R. and Namboodiri, N.R. (1964). A Test of the Accuracy of the Ratio Correlation Population Estimates. Land Economics 40: 100-102.

[4] Government of British Columbia Central Statistics Bureau (1980). British Columbia School District Population Estimates, and also, British Columbia Municipal Population Estimates. Victoria, Ministry of Industry and Small Business Development, Government of British Columbia; unpublished report.

[5] Johnson, T. (1963). Econometric Methods, New York, McGraw-Hill Book Company, p. 219.

[6] Mandell, M. and Tayman, J. (1982). Measuring Temporal Stability in the Regression Models of Population Estimation. Demography, 19(1): 135-146.

[7] National Research Council (1980). Estimating Population and Income of Small Areas. Washington, D.C., National Academy Press. pp. 1-247.

[8] Norris, Douglas A., Britton, Malcolm and Verma, Ravi (1982). The Use of Administrative Records for Estimating Migration and Population. Presented at the Annual Meeting of the American Statistical Association, Cincinnati, Ohio, August 16-19.

[9] Norris, Douglas A. and Standish, Linda D. (1983). A Technical Report on the Development of Migration Data from Taxation Records. Ottawa, Statistics Canada. Draft.

[10] O'Hare, W. (1976). Report on a Multiple Regression Method for Making Population Estimates. Demography 13: 369-380.

[11] O'Hare, W. (1980). A Note on the Use of Regression Methods in Population Estimates. Demography 7: 87-92.

[12] Raby, R. and Parent, P. (1982). Postcensal Emigration Estimates 1980-1982. Ottawa, Statistics Canada, (Mimeographed). pp 1-29.

[13] Romaniuc, A., Raby, R. and Parent, P. (1982). The Choice of Methods for Estimating Interprovincial Migration for the Post-1981 Period. Ottawa, Statistics Canada, (Mimeographed). pp. 1-43.

[14] Spar, M. and Martin, J. (1979). Refinements to Regression-Based Estimates of Postcensal Population Characteristics. Review of Public Data Use 7 No. 5/6.

[15] Statistics Canada, Annual. Estimates of Population for Census Divisions. Cat. 91-206. Ottawa, Ministry of Supply and Services, Government of Canada.

[16] Statistics Canada, Annual. Estimates of Population for Census Metropolitan Areas of Canada. Cat. 91-207. Ottawa, Ministry of Supply and Services, Government of Canada.

[17] Swanson, D. (1978). An evaluation of Ratio and Difference Regression Methods for Estimating Small, Highly Concentrated Population: The Case of Ethnic Groups. Review of Public Data Use 6: 18-27.

[18] Verma, Ravi B.P. and Basavarajappa, K.G. (1982a). A Sub-provincial Estimation of the Population in Canada: A review of Estimation Methods and Prospects for Development. Presented at the Population and Family Planning Session, American Public Health Association Meeting, Montreal, November 1982. pp. 1-32.

[19] Verma, Ravi B.P., Basavarajappa, K.G. and Bender, Rosemary (1982b). New Approaches to Methods of Estimating the Population of Census Divisions. Ottawa, Statistics Canada, (Mimeographed). pp. 1-77.

[20] Verma, Ravi B.P., Basavarajappa, K.G. and Bender, Rosemary (1982c). New Approaches to Methods of Estimating the Population of Census Metropolitan Areas. Ottawa, Statistics Canada, (Mimeographed). pp. 1-39.