Use of Administrative Data Files for Migration Estimates: A Case Study of Driver's Licence File in Ontario[1]

RAGHUBAR D. SHARMA and CHEUK WONG[2]

ABSTRACT

In Canada, provincial and federal demographers have attempted to use various sets of administrative data to estimate migration flows. This paper presents the development of intra-provincial migration estimates using driver's licence data in Ontario. An evaluation of these migration estimates has been carried out by comparing with those derived from the income tax data by Statistics Canada. Both files provide equally good and complimentary estimates of intra-provincial migration.

KEY WORDS: Administrative files; Population estimates; Component method; Small areas; Error of closure; Intraprovincial migration.

1. INTRODUCTION

Migration is an important component of population projections, and population estimates. As no records regarding the movement of population are kept in Canada, demographers in the federal and provincial governments have attempted to use various sets of administration data to estimate migration flows. Statistics Canada uses revenue data (Norris and Standish 1983), British Columbia utilizes hydro-hookups (McRae 1985), and Alberta uses health care records (Alberta Bureau Statistics 1985). Since 1979, Ontario has been using drivers' licence address changes to estimate intra-provincial migration. Apart from the quality aspect, one major attractiveness of the driver licence data is in its timeliness. There is only a 4 to 5 week time lapse between receiving the data and the date of reference compared with over one and one-half years in revenue data. In this paper we shall present an evaluation of estimates of intra-provincial migration derived from the driver's licence data in Ontario. In the U.S.A., the State of California also uses driver's licence address changes for the estimation of intra-provincial migration (Hoag 1984).

2. DRIVER'S LICENCE DATA FILE

Information on driver's licence address changes is made available by the Ontario Ministry of Transportation and Communications (MTC). A driver is required to notify the Ontario Ministry of Transportation and Communications within 90 days of his/her change of address. The information is available at the postal code area level. These postal code areas can be converted into such subprovincial areas as, counties, regions and municipalities. As Table 1 indicates, data are available for the past seven years. Since 1979, data are also available for each quarter of these years.

More than a million changes of addresses are recorded every year. The majority of these moves tend to be within census divisions (that is, county or regional municipality). However, net inter-county movers averaged only about 22,000 per year. Table 1 indicates that about one-third of the records do not provide a postal code for either origin and/or destination of the mover.

In Ontario, a person becomes eligible to hold a driver's licence at the age of 16 years. More than 75 per cent of the eligible population holds a driver's licence. The elderly population and female population have a much lower tendency to hold a drivers' licence (Table 2).

3. CONVERSION OF DRIVERS TO MIGRANTS

An adjustment factor is applied to the number of drivers to arrive at the number of movers. This adjustment factor (F) is calculated as follows:

$$FA = \frac{\text{Known and Unknown Movements}}{\text{Known Movements}}$$

$$FB = \frac{\text{Total Population}}{\text{Population with a Licence}}$$

$$F = FA \times FB.$$

Table 1

Number of Total Movers and Number of Movers with
Unstated Origin and/or Destination, Ontario, 1975-1985

Year	No. of Known Movers (Inter & Intra Country)	No. of Origin and/or Destination Unstated	Total	% Unstated
1979 (Calendar Year)	881,000	0	881,000	0
1979/80	586,000	301,000	887,000	34
1980/81	566,000	306,000	872,000	35
1981/82	617,000	270,000	887,000	30
1982/83	648,000	259,000	907,000	29
1983/84	822,000	320,000	1,142,000	28
1984/85	831,000	330,000	1,161,000	28

Source: Ontario Ministry of Transportation and Communications.

Table 2

Percent of Population Holding Driver's Licence, Ontario, 1981

Age	% of Population Holding A Driver's Licence, 1981		
	Male	Female	Total
16-19	63.9	36.5	49.1
20-24	92.7	73.0	85.7
25-34	98.6	81.6	90.0
35-44	99.7	79.9	90.0
45-54	96.7	67.8	82.4
55-64	93.2	56.7	74.2
65 +	73.1	27.4	46.4
Total	90.6	62.2	75.8

Source: Ontario Ministry of Transport and Communications.

FA accounts for the unstated origins and/or destinations and *FB* accounts for non-driver's licence holders. The factor assumes that migration patterns of those who do not hold driver's licence do not differ from those who hold driver's licence. Similarly, it assumes that migration patterns of those with unstated movements do not differ from those whose movements are stated.

4. INTRA-PROVINCIAL MIGRATION ESTIMATES: DRIVER'S LICENCE VERSUS INCOME TAX FILES

Statistics Canada uses change of address as provided by a taxpayer on his annual income tax return. The number of children are estimated from the number of dependents claimed by the taxpayer. Like the driver's licence data, adjustment factors have to be introduced to the revenue data to overcome unstated postal code and people who do not file an income tax. Furthermore some taxpayers use a non-residential mailing address in their return.

The relative accuracy of migration estimates derived from the income tax file and driver's licence file needs to be tested. Three measures have been applied to test this relative accuracy of the two data sets. They are:

A. Errors of Closure
B. Growth Rates Test
C. Index of Dissimilarity

Ideally, errors of closure and growth rates should be calculated from the population estimates from one census year to the next. Reliable data on driver's licence address changes are available only from 1979 onwards in Ontario. Therefore, 1979 intercensal population estimates of Statistics Canada and estimated 1981 population were used as base. Two sets of population estimates were calculated. First, using the driver's licence address file for intra-provincial migration and second, using the income tax file for intra-provincial migration. All other components, i.e., births, deaths, interprovincial migration and international migration were kept the same for both sets of population estimates.

4.1 Errors of Closure

Two sets of population estimates for the census divisions (one using driver's licence data and the second, using income tax data for intra-provincial migration) were compared with the 1981 census population. The percent difference in the estimated population from the census population is called *error of closure*. Out of 49 census divisions, 23 have smaller errors if driver's licence data are used and 26 census divisions have smaller errors if income tax data are used to estimate intra-provincial migration.

4.2 Toronto Urban Complex

A quite interesting picture emerges in the Toronto Urban complex which includes six regional municipalities (Table 3). Driver's licence data yields a smaller error of closure for the complex as a whole and under-estimates the population for the areas outside of Metro Toronto.

The income tax file gives lower errors of closure for individual census divisions within the complex whereas, for the complex as a whole the error is larger than the driver's licence file (Table 3). Accordingly, driver licence data were used for estimating intra-provincial migration for the Toronto complex as a whole and the distribution to individual census divisions of the complex was based on revenue data.

4.3 Population Growth Rates

Percent change in the population estimates from 1979 to 1981 was calculated for the estimates derived by using the driver's licence file and the income tax file respectively. These growth rates were compared with the 1981 census growth rates.

Table 3
Errors of Closure for Toronto Urban Complex

Census Division	Errors of Closure	
	Income Tax	Driver's Licence
Durham R.M.	−0.21	−0.41
Halton R.M.	0.45	−0.51
Hamilton-Wentworth R.M.	0.10	−0.11
Peel R.M.	−0.63	−1.94
Toronto R.M.	0.01	1.26
York R.M.	−0.11	−5.70
Total Toronto Urban Complex	−0.05	−0.01

There is not much difference in the relative closeness of growth rates of the two sets to census growth rates. The number of census divisions which yield different direction of population change than the census growth rate are 3 for the driver's licence data and 10 for the income tax file (Table 4). This is one aspect where the driver's licence data appeared to yield more reliable estimates than the revenue data.

4.4 Index of Dissimilarity

Index of dissimilarity was calculated for in- and out-migration separately, as the direction of net migration was not the same for some counties for the two sets of estimates. The value of the index of dissimilarity can vary between 0 and 100. It is the half of the sum of the absolute differences between the two corresponding percent distributions and is equivalent to the sum of the positive differences or the sum of the negative differences (Shryock and Siegel 1971). The general formula is:

$$ID = \tfrac{1}{2} \sum |r_2 - r_1|$$

where, r_2 and r_1 are the corresponding percentages in the two distributions.

The low values of the index indicate that both files (the driver's licence and the income tax) yield quite similar estimates of intra-provincial migration for the census divisions of Ontario. However, over the four years the extent of dissimilarity increases for out-migration and improves for in-migration (Table 5).

5. CONCLUSION AND SUMMARY

This study attempts to compare the intra-provincial migration estimates derived from the driver's licence file with those derived from the income tax file. Both files provide reasonably good measures of the magnitude of intra-provincial migration for the Census Divisions of Ontario.

Although the driver's licence data appeared to provide better estimates in the direction of intra-provincial migration, the income tax data resulted in slightly more counties with smaller errors of closure and in addition yielded somewhat better results in some major areas (for example, distribution within the Toronto/Hamilton urban complex). In view of their respective strengths, the appropriate approach is to combine the use of these two data sources.

Another issue that should be noted is that the evaluation was based on three years only i.e., 1979 to 1981. A more accurate assessment on the quality of these two data files cannot be made until the availability of the 1986 census data.

To further improve the quality and the applications of the driver's licence data, the following two areas are suggested for further research:

- Verification of *FA* factor based on actual counts of unknown origin/destination through using manual coding of addresses.
- Extension of the use of the driver's licence data file as an additional source to family allowance and revenue data for inter-provincial migration estimates.

The driver's licence file tends to over-estimate migrants for Metro Toronto and under-estimate for the areas surrounding Metro Toronto. The reverse seems true for the income tax file. For this region as a whole, the driver's licence file gives better estimates for intra-provincial migration than the income tax file. The income tax file provides a better distribution of intra-provincial migrants in the counties of this region.

Table 4
Census Divisions Which Yield Different Direction of
Population Change Than Those Based on Census

Census Division	% Changed Based On	
	Income Tax	Census
Bruce	0.11	−1.77
Grey	−0.04	0.17
Hastings	0.12	−1.03
Leeds and Grenville	0.21	−0.32
Niagara	0.15	−0.46
Northumberland	0.31	−0.82
Oxford	0.17	−0.09
Parry Sound	1.20	−0.51
Stormond/Dundas/Glengarry	0.44	−0.17
Sudbury T.D.	0.41	−2.28
Province	1.60	1.46

	% Change Based On	
	Driver's Licence	Census
Leeds and Grenville	0.02	−0.32
Parry Sound	0.81	−0.51
Thunder Bay	0.42	−0.20
Province	1.60	1.46

Table 5
Index of Dissimilarity

Year	Index of Dissimilarity	
	In-Migration	Out-Migration
1979-80	5.61	3.50
1980-81	5.54	3.82
1981-82	5.26	4.82
1982-83	4.41	4.87

REFERENCES

ALBERTA BUREAU OF STATISTICS (1985). The development of Alberta health care records and their application to small area population estimates. A paper presented at the Meetings of the Federal-Provincial Committee on Demography, Ottawa.

HOAG, ELIZABETH (1984). Estimating annual migration for California counties using driver's licence address change. A paper presented at the Meetings of the Population Association of America, Minneapolis, Minnesota.

McRAE, DONALD G. (1985). The use of hydro accounts in the British Columbia based population estimates. A paper presented at the Meetings of the Federal-Provincial Committee on Demography.

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

SHRYOCK, HENRY M., and SIEGEL, JACOB S. (1971). *The Methods and Materials of Demography*. Washington: U.S. Bureau of Census.