Canadian Internal Migration Statistics: Some Comparisons and Evaluations*

John Vanderkamp and E. Kenneth Grant
Department of Economics, University of Guelph,
Guelph, Ontario N1G 2W1.

Introduction

Migration statistics play an important role in demography, in labour force planning, and in empirical research for testing analytical propositions. Migration statistics are obtained from a variety of survey sources for different sample sizes, relate to different population bases, refer to different time periods and to different geographical units. As a result, the patterns revealed by the various migration data are by no means the same. The choice of data for a particular research purpose is therefore not a matter of indifference.

A demographer interested in estimating regional populations will likely desire migration statistics about large random samples of the whole population, with possibly some age breakdowns. An economist who wants to apply the human capital model to regional adjustment in the labour market will probably require migration data for the working-age population, preferably with a directly related data base of relevant labour market variables such as wages, unemployment, and education. A regional scientist, or a sociologist, examining the role of earlier migration experience on current migration decisions, may prefer micro data of a longitudinal type. A geographer who is studying the regional demands for services to the elderly (such as nursing and retirement homes) may be particularly interested in the migration

* Original paper given at the Migration Sessions, jointly organized by the Canadian Regional Science Association and the Canadian Association of Geographers, held at McMaster University on 25-30 May 1987. Comments by Jacques Ledent, Neil Field, and an anonymous referee were very helpful in the process of revising this paper. We are grateful to our research assistants Mike Kupferschmidt and Reagan Pratt for their expert help.

ISSN: 0705-4580 Printed in Canada/Imprimé au Canada

behaviour of people who are retiring. These examples are designed to illustrate the obvious proposition that what is desired as the migration data base will depend on the purpose and framework of the researcher.

But the real world is often not the most desired state of affairs. We may only have one or two sets of migration data, which may have to serve all these and other research purposes. Our demographer and our economist may both have to contend with migration data derived from family allowance transfers, particularly if the economist is interested in cyclical variance in behaviour. Our regional scientist may at best have access to annual family allowance data and census data on 5-year migration in order to make some inferences about repeat migration patterns. Our geographer may have to be content with an age breakdown of census data of 5-year migration. In each of these cases the results obtained by the researcher will in part be determined by the choice of migration data, and in some cases the results may be wrong or the inferences may be inappropriate.[1] Some researchers may, of course, be able to conduct their own migration surveys for their own particular purposes, but this represents an expensive route, and we concentrate on the evaluation of existing data bases.

It is not the purpose of this paper to indicate which data set is appropriate for which study of migration. That is the prerogative and the responsibility of each individual researcher. Our much more limited objective is to discuss some of the major sources of existing internal migration data, especially interprovincial migration data, and to provide some comparisons and evaluations of the various data sets. The analysis will relate to gross migration rates, correlations of migration matrices, net-gross migration ratios, and net impact measures, and it will also include specific case studies of the migration experience of some Canadian provinces in the last few decades.

Data Sources

In describing the various migration statistics we shall concentrate on broad principles. Each statistical data collector needs to make decisions about the concept of migration and how it is to be measured, and about coverage in terms of people, time and place. In addition, some other features of a data base will be mentioned, particularly those related to linkage with other data.

[1]To give only one example, Winer and Gauthier (1982) use both family allowance and income tax data on migration to test a variety of hypotheses, particularly about fiscally induced migration behaviour, and the results for the two data sets are by no means identical. Other examples of data usage and its inevitable compromises are contained in the list of references.

Statistics Canada has recently published a very useful monograph *Population Estimation Methods, Canada* (Statistics Canada 1987), which describes in some detail the internal and external migration data being used and generated by Statistics Canada at present (in Chapters 4 and 5, respectively); references to this publication will permit us to be relatively brief in our description of two of the main data series on migration.

Family Allowance Data (FAMALL)

The family allowance system is administered by provincial branches within the federal Department of Health and Welfare. These branches record the change-of-address notices from families receiving family allowances. Estimates of interprovincial migration based on family allowance data were published by Statistics Canada until 1981 (cat. 91-208). They are available in published form back to 1961, but the basic transfer data can be used back to 1947 (Vanderkamp 1968); the original statistics were evaluated by Kasahara (1963). The transfers are recorded on a monthly basis, but they are generally only used as annual sums.

The decision to submit a change of address is motivated by a change in the province of residence that is thought to be relatively permanent. A person making a move that is known beforehand to be temporary may or may not submit a request for a transfer of account. This scenario implies that the time span for the migration decision is not known and is dependent on individual judgments. Some return and repeat migration decisions within a calendar year are recorded in the statistics (Statistics Canada 1987, 60-63).

To compute a migration rate we need to know what the appropriate denominator is; that is, we need to know the "population" at risk. Since these migration statistics are based on the continuous recording of transfers, it is in principle impossible to determine the appropriate population at risk. Moreover, over the last four decades changes in the eligibility rules of children and families, and changes in the average family size, have affected the link between this data set and migration measurement for the population at large. The best route may indeed be simply to use the average annual population of eligible families (see Vanderkamp 1968) or children (Statistics Canada 1987, Ch. 4).

This question of the link between the migration patterns of children and adults takes up a major part of the migration chapter in *Population Estimation,* as this issue is crucial for the development of migration estimates (and thus population estimates) for the population as a whole. The multiplicative factors used by Statistics Canada for this purpose have seen an evolution over time towards increased sophisti-

cation, and since 1981 they are derived from the taxation statistics on migration, to be discussed next.

In short, while these family allowance statistics of migration flows have some conceptual difficulties, they provide the longest time series of provincial migration matrices and they are available with a very short lag. Two migration series based on these data were published on a 12-months basis (June to May) back to 1961-2, one for eligible families and the other for total population. Statistics Canada discontinued publication of interprovincial migration estimates based on FAMALL data in 1981, although they continue to be used to generate preliminary post-censal population estimates. For final post-censal estimates of population, Statistics Canada now uses the migration data derived from income tax sources, because they are presumed to be superior to the FAMALL data. Winer and Gauthier (1982) present versions of both series on an annual (calendar year) basis for the period 1950-78.

Income Tax Migration Data (INCTAX)

Once personal identification numbers, such as Social Insurance Numbers, are introduced into a set of administrative records, the study of longitudinal phenomena becomes an obvious possibility. Providing the records contain information on individual locations, a study of migration can be undertaken. A micro data base of this kind, even derived from a sizeable sample, is bound to run into confidentiality problems, made all the more severe by the availability of reported income figures (as when the income tax system provides one of the main sources of information). Administrative units are not set up to provide public use sample tapes, with the result that such data bases are not generally in the public domain; moreover, users may include other government departments, which are inclined to work with all available detail. This poses a problem for researchers, who typically have to exploit established contacts to develop possibly complicated contracts to set the framework within which the research may be conducted. As an alternative, the data may be released in tabulated form, rather than the much richer micro form (Winer and Gauthier 1982). Statistics Canada now releases tabulated migration data derived from INCTAX sources (cat. 91-210).

The income tax source contains three possible pieces of information on an individual's location. The first two are obtained from the so-called T1-form completed by the filer at the time of tax submission. The T1-form relates the province of residence on 31 December of the relevant tax year, labelled as "taxing province" for our purposes. The same form also has the mailing address of the individual at the time of mailing the tax form (for most taxfilers, in March-April of the year following the tax year); this may be combined into provincial aggregates ("mailing province"), or sub-provincial regions, or localities (Grant and Vanderkamp 1976). By linking T1 data for different years, one obtains the basic migration data; the time span is about one year and the timing of migration decisions is fairly clear. This type of income tax information is exploited by Statistics Canada (1987, 54-59) in generating their migration estimates referred to above. The third piece of location information can be derived from the so-called T4-slips, which are completed by employers to indicate the amount of employment income earned (as well as various deductions) by an individual during a tax year. These T4-slips also contain information on the person's province of employment, and again linkage for different years provides the basic migration data. A person may, of course, have T4-slips indicating different provinces of employment in a particular tax year, and this leads to some ambiguity about the time span and the timing of migration decisions.

Income tax information is incorporated into two micro data bases: the "old" Unemployment Insurance (UI) data base covering the years 1965-71 and the "new" UI data base covering the period since 1972; unfortunately, the two UI data bases cannot be linked. The old UI data base was put together by the UI Commission to study the impact of possible UI policy reform, a reform that saw the light of day in the 1971 UI legislation. It uses a 2 percent sample of all individuals issued with Social Insurance Numbers (this SIN source also gives age and sex), and aside from income tax data from T1 forms only (which also includes information on different sources of income, dependency, occupation, industry, and union status) it incorporates UI data on premium contributions, claims and benefits. We have made intensive use of this data base both in tabulated form (Grant and Vanderkamp 1976) and in its micro data form (Grant and Vanderkamp 1980; 1986). The coverage of these migration data relates to income tax filers, which is, of course, not the same as coverage of total population; for researchers primarily interested in the adjustment behaviour of labour market participants, this limited coverage is not a real handicap.

The new UI data base has been used in the analysis of labour turnover, such as quits and lay-offs, and labour market adjustment. It is available in different sample sizes, such as 10 percent, 1 percent and 0.1 percent of SIN cases, for the period since 1972. The new UI data base contains other information that is similar to that in the old UI data base, but the new one has more detailed information on labour market changes, including changes of employer, and much less detailed data from the income tax source, since only the T4 forms are included and no T1 information. As already indicated, the migration data from this source are derived from the T4 slip information about province of employment. Since we have not worked much with the migration data

of this source, for the comparison exercise below we have adopted a somewhat ad hoc procedure for dealing with the cases of multiple provinces recorded within a year; that is, we have simply designated as the province the one indicated on the T4 slip reporting the largest earnings level. This means that there is uncertainty about the time span used and about the precise timing of migration decisions, although we assume that the migration decision was made during the 12-month period July to June. In short, while there are similarities between the two INCTAX data bases, the two migration series must be considered as separate data sets, since concepts, coverage, and time span are all different.[2]

Tabulated income tax filer data reported by Winer and Gauthier (1982) have already been referred to above; they are obtained from a 10 percent sample of taxfilers, and they relate to changes in taxing province during the calendar year. Statistics Canada has published tabulated INCTAX migration data (based on a 10 percent sample, and using the mailing address information) since 1983 (cat. 91-210); the data that we obtained from Statistics Canada cover the years 1976-85.[3] There was also a migration data set based on income tax data for 1963-64, which is referred to in Vanderkamp (1968) but about which little more is remembered.

Census Migration (CENS)

This survey of migration is a component of the "long" census form distributed to a random sample of the population at the time of the census taking. The migration data are derived from a question about the individual's residence location five years prior to the census date; hence we use the term 5-year retrospective migration. Migration data for five different census dates are available in published form—1961, 1971, 1976, 1981, and 1986; the 1986 data are only just being released. Sample sizes were 20 percent in 1961, 1981, and 1986, while in 1971 and 1978 a one-third sample was used. Canadians serving outside Canada are excluded from the samples.

Because of the retrospective nature of the survey the migration data are only relevant and reported for the population 5 years and over; migration for the 0-4 year age group can be inferred with reference to place of birth. Immigrants who arrived in Canada in the 5-year

[2]The old and new UI data bases are not available to the "public", and they are not easily obtained by researchers either. Researchers need to establish special contacts (and contracts) with the Canada Employment and Immigration Department to obtain access, and even then the research may have to be conducted "in-house" at the department.

[3]INCTAX data were also compiled by Statistics Canada for the period 1966-75, but these data are considered to be "experimental" and are not readily made available.

period are identified separately and can therefore be eliminated for the calculation of internal migration rates.

The sample sizes of the CENS data base are very large, and a number of geographical disaggregations are possible, including provinces, major urban areas and census divisions (equivalent to counties in some provinces), and potentially to the census subdivision level. The CENS migration data provide direct linkage with the large information base of the census, including such variables as sex, age, marital status, education, ethnic origin, mother tongue, home language, employment status, and income level. Place of birth and education have also been used to study long-term migration patterns, particularly in relation to return and repeat migration. Not all the cross-classifications based on these variables are available in published form, and a detailed request for data can cost many thousands of dollars. Detailed cross-classifications are also likely to lend to confidentiality problems, with small cells being suppressed.

Although we have not worked with the so-called Public Use Sample micro data bases of the 1971, 1976, and 1981 Censuses, some researchers have made good use of them (Marr and Millerd 1980; Robinson and Tomes 1982). While they relate to a much smaller sample (2 percent in 1981, and 1 percent in 1971 and 1976) of the relevant populations, the same direct link with the whole census information base is in principle available. The only problem is that due to sample size and confidentiality restrictions these micro data only relate to interprovincial migration in terms of place-to-place flows (Liaw and Ledent 1988); other types of geographical movers can be identified, such as residential, intermunicipal, and intercounty movers, but no place-to-place information is available at these levels.

Insured Population Migration (UIPOP)

This source of data is only of historical interest, since the change in the operation of the unemployment insurance system in 1971 has in effect discontinued the routines on which this data series was based. Every year on June 1 a member of the Insured Population had to renew his or her UI book; the required information was supplied by the employer or the UI office and included age, sex, occupation, industry, province of employment, and claimant status. The Insured Population includes all persons eligible for unemployment insurance, which covered about 75 percent of wage earners in 1961 and about 80 percent in 1968. By using the individual's insurance number (later SIN), data from different years are linked.

Greenway and Wheatley (1963) report mobility data for the period 1952-59 based on a 1 percent sample of the Insured Population. Vanderkamp (1973) reports data based on a 10 percent sample for the

years 1965-68. The strength of this data source was that it provided information on occupational and industrial mobility alongside migration; unfortunately the information on occupational changes was unreliable. The time span was obviously one year and migration decisions were timed in the 12-month period from 1 June. The migration concept relates to changes in place of employment and its coverage is unique, which makes the UIPOP data less readily comparable to the other available data.[4]

Net Changes Due to Migration (SURV)

The abbreviated name of this data series (SURV) is based on the demographer's term, "survival" method, typically applied to this approach. The method consists of subtracting "natural increase" from population change for regions between two time points; for total population natural increase is simply births minus deaths in the relevant period, while for the working-age population (15+) it is the 14-year olds in the preceding year minus deaths in the 14+ population.

The SURV data have a number of disadvantages. First, the migration data are not available in matrix form but only as net changes due to migration. Second, the method inevitably includes international migration with domestic migration; if separate data are available on net immigration by region, then estimated net changes due to internal migration can be derived. If one is interested in regional labour supply effects in their entirety, this limitation may be considered an advantage. Third, the survival method presumes that the data on population and "natural increase" (and net immigration) are reliable, since any errors in these data will tend to be reflected in the residual calculation of (internal) migration. Fourth, SURV data based on post-censal estimates of population may contain "forecasting" errors associated with these estimates. But the method has one major advantage: it can readily be applied in a consistent way to different time periods and regions, and thus it provides the only migration data for a considerable period of our history.

As already indicated SURV data on migration can be derived for total population and for different subgroups. If the population data apply to 1 June of each year and if the natural increase data are aligned, then the implicit time span is one year and the migration decision's timing is in the 12-month period from 1 June. As population counts by region refer to place of residence, so does migration refer to changes of residence. The SURV data used in the section on "Impacts of Migration", below, were originally calculated for some recent research on regional adjustment (Vanderkamp 1988a; 1988b).

[4]These data are probably no longer available, in any form because of their old age. The interprovincial migration matrices for the three years 1965-6, 1966-7 and 1967-8 are published in Vanderkamp (1973, Appendix E), which also contains return migration matrices for the last two years.

Other Migration Statistics

Under this heading we simply mention data sets that we happen to know about but that are more sporadic in nature. From time to time the Labour Force Survey (LFS) is used to examine a particular topic in detail, and migration has been looked at twice in this context. Nickson (1967) deals with the results regarding migration patterns for the 12-month period ending October 1965. In 1981 Statistics Canada undertook a special survey regarding migration during the period 1976-80, also as part of LFS (Statistics Canada 1982; 1983), which has been utilized for research purposes by Marr and Millerd (1987; 1988).

Mobility Rates, Correlations, and Net-Gross Ratio

Before we report the summary statistics for the migration data to which we currently have access, we will briefly review the evidence on migration rates and correlations given in earlier work. In Vanderkamp (1968, 599) it is reported that there is a very strong similarity between the FAMALL data for 1964 (with seven regions, including Atlantic) and income tax filer data on migration for the same year, with a correlation coefficient of 0.98. The same paper also indicates that the net migration statistics from FAMALL and (male labour force members) from the 1961 Census for the period June 1956 to June 1961 are quite similar.

In Vanderkamp (1973, 14-16) the UIPOP migration data are compared with the three data series just mentioned, and also with the Labour Force Survey data (LFS) reported in Nickson (1967). The calculated overall migration rates (with 10 provinces) were as follows:

UIPOP 1965-6 (June)	2.7%
UIPOP 1966-7	2.8%
UIPOP 1967-8	2.4%
Income tax filers 1964	1.5%
FAMALL 1965-6 (June)	2.0%
LFS Pop.14+ 1964-5 (October)	1.3%
Census labour force 1956-61 (June)	3.8%

These migration rates are computed by dividing the sum of the elements in the migration matrix (excluding stayers in the diagonal) by the relevant sample population at risk (that is, with records at both time points). The migration rates derived from UIPOP data are on the

high side compared with the other rates; it should be recalled that these refer to the insured population and to changes in the province of employment. The rate for the census data refers to the 5-year migration time span, and this rate is probably consistent with 1-year migration rates of about 1.5 percent. We shall compare 1-year and 5-year rates below. Only two pairs of migration data refer to roughly the same periods, so correlation makes some sense in these cases. The UIPOP 1965-6 and FAMALL 1965-6 migration matrices (all elements are "normalized" by the populations in sending provinces) have a correlation coefficient of 0.71, and the corresponding correlation coefficient for income tax filers 1964 and LFS 1964-5 is 0.83. These are not particularly high levels of correlation, and we shall see whether this kind of pattern holds up in the more detailed and more complete comparisons to which we now turn.

Figure 1 shows the time patterns of migration rates computed for a number of the series derived from three basic data sources: family allowance, income tax, and census. The pictured time series are clearly identified at the bottom of Figure 1. There are migration rates from a total of seven different data sets, with considerable overlap in time periods:

(i) FAMALL (STATCAN) from Statistics Canada 91-208, relating to families for 12-month periods from June 1961-2 to 1985-6 with 10 provinces + North;
(ii) FAMALL (W&G) from Winer and Gauthier (1982, Supplement) relating to families for calendar years 1950 to 1978, with 10 provinces.
(iii) INCTAX (OLD UI) from the old UI data base, based on mailing province, for 12-month periods from April, 1966 to 1972, with 10 provinces + North; the data for 1967-8 and 1968-9 contain errors and are therefore omitted from Figure 1;
(iv) INCTAX(W&G) from Winer and Gauthier (1982, Supplement), based on taxing province for calendar years 1968 to 1977, with 10 provinces;
(v) INCTAX (NEW UI) from new UI data base, based on T4 information related to province of employment, for roughly 12-months periods from about July, 1972-3 to 1982-3 with 10 provinces + North;
(vi) INCTAX (STATCAN) from Statistics Canada 91-210, based on mailing province, for 12-month periods from April, 1976-7 to 1985-6, with 10 provinces + North;
(vii) CENSUS from 1961, 1971, 1976, 1981, and 1986 Census sources, covering the five-year periods preceding census dates, with 10 provinces.

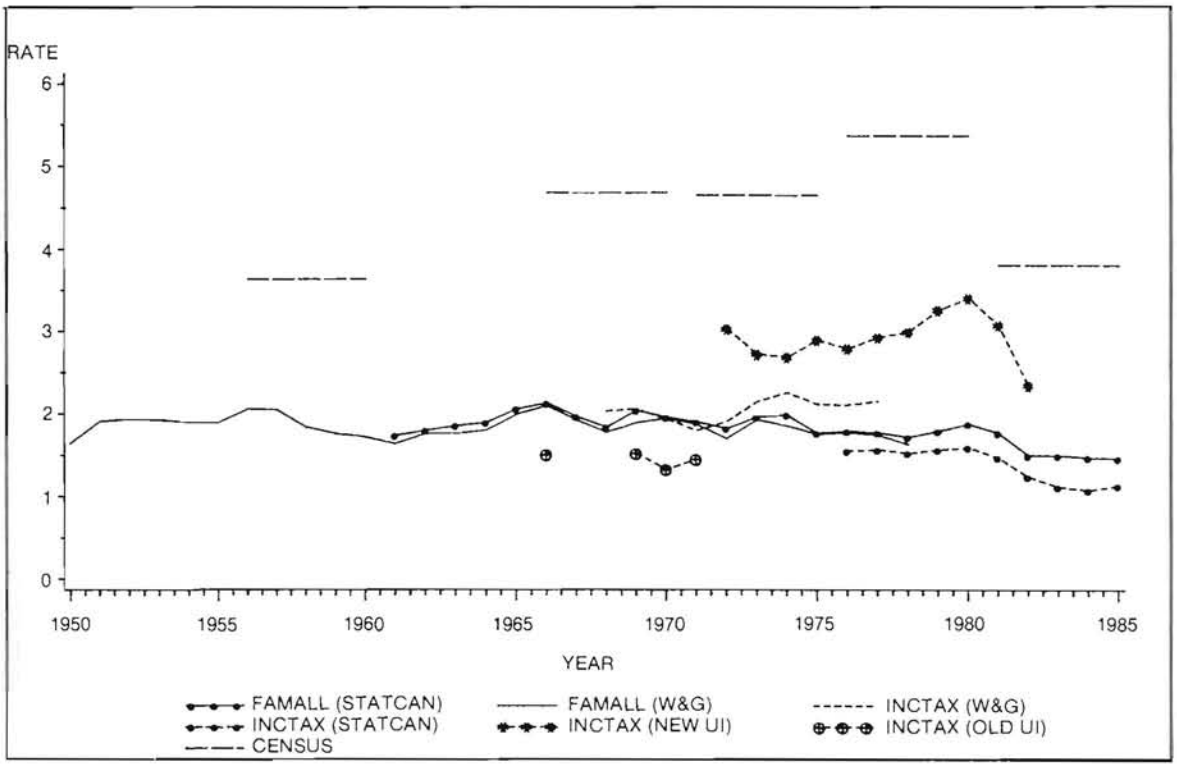

Figure 1

INTERPROVINCIAL MIGRATION RATES
(MIGRANTS EXPRESSED AS PERCENTAGE OF BASE YEAR POPULATION)

There are substantial differences between the four sets of INCTAX annual migration rates, with (iii) at about 1.5 percent, (iv) at about 2 percent, (v) at close to 3 percent, and (vi) at about 1.5 percent. As already indicated, we are not very familiar with the migration data of (v), and we can only explain its high migration rates as being in part related to multiple provinces of employment arising from more than one T-4 slip for a significant number of people who may not have changed their permanent residence. The main difference between the migration rates of (iii) and (iv) is related to the fact that (iii) is based on a subset of individuals with valid tax records in each of the years 1968 to 1971. By eliminating individuals who became or stopped being tax filers (or were intermittent filers), we eliminated a disproportionate number of interprovincial migrants. This subset became our micro data base for (Grant and Vanderkamp 1980); the original data displayed migration rates of about 2 percent (Grant and Vanderkamp 1976, 14). In short, if we discount the data from (v), it appears that migration rates from income tax sources are roughly of the same order of magnitude as those obtained from family allowance sources, but the FAMALL migration rate declined somewhat less in the 1980s than the INCTAX migration rates.[5]

The nature of the INCTAX micro data bases allows us also to calculate migration rates with a 5-year time span, for comparison with the corresponding annual rates and with the 5-year rates from the CENS data; these INCTAX 5-year rates are not shown in Figure 1. The 5-year rate calculated from the old UI data base for the period April 1966 to April 1971 is 4.3 percent (compared with 4.7 percent for the 1966-71 CENS data), and the rate for the new UI data base for the period July 1976 to July 1981 is 8.1 percent (compared with 5.4 percent for the 1976-81 CENS data). It is obvious that the 5-year migration rates are much less than five times the average annual rate; a factor of roughly 2.5 to 3 appears more appropriate. Probably the most important explanation for this difference is the phenomenon of repeat migration, including return moves. If about one-third of all interprovincial migrants are repeaters and about half of the repeaters are return migrants, then a 5-year migration rate would be about three times the average annual rate; such rates of repeat migration are probably not out of line with reality (Grant and Vanderkamp 1984; 1986). In addition to repeat migration there is the problem of sample turnover, which means that, in general, each of the 1-year migration

[5]Jacques Ledent has suggested that this pattern is not unexpected, since the 1982-3 recession may have had a relatively more serious impact on wage earners, who are more important among INCTAX samples, than on families with young children, who make up the FAMALL sample.

rates is computed for a somewhat different set of persons and the 5-year rate is computed for a subset of all these persons.[6]

The CENS migration rates increased from 3.6 percent for 1956-61 to 5.4 percent for 1976-81, to decline again to 3.8 percent during the 1981-6 period. This presents us with a bit of a puzzle: why is there such a strong upward trend in the 5-year migration rates, obtained from the CENS source, during the period 1956 to 1981 when there is little or no trend in the annual migration rates derived from other sources and in particular from the FAMALL source? Two possible explanations come to mind. First, the incidence of repeat migration may have diminished over this period; but we have no data to support this. Second, the migration propensities of other members of the population (other than families), and particularly of unattached individuals, may have changed over time. The increase of young people entering the labour market during the seventies and early eighties is at least consistent with this. On the other hand, the decline in 5-year CENS migration rates in the eighties is roughly consistent with the pattern of the other series in Figure 1.

Our general conclusion is that migration rates may vary considerably depending on definitions and samples. Moreover, there is a significant amount of time series variation, with evidence that the migration rate varies pro-cyclically (Vanderkamp 1968). The 1982-3 recession appears to have reduced migration rates quite drastically, although the decline actually started in 1980.

Two migration matrices may have a similar migration rate but they may not display the same interregional pattern. Correlation analysis of all the elements in two migration matrices is a useful tool for comparisons between interregional patterns for the same time periods.

[6]Unfortunately, we can only check this sample turnover problem with our INCTAX (NEW UI) data base. The following table shows the sample sizes and migration rates plotted in Figure 1, and also the migration rates obtained from the restricted sample (size 7316) with the restriction that persons must have a record in each of the years 1976-7 to 1980-1.

	Original samples		Restricted sample(7318)
	Size	Rate	Rate
1976-7	9515	2.8%	2.8%
1977-8	9620	2.9%	2.7%
1978-9	9870	3.0%	2.6%
1979-80	10273	3.3%	3.0%
1980-1	10592	3.4%	2.9%
1976-81	8305	8.1%	7.6%

Obviously, sample sizes vary a good deal and the restriction in sample selection lowers migration rates. On the other hand, the restriction has very little effect on the *ratio* of the 5-year migration rate to the average annual rate, which only changes from 2.6 to 2.7.

As expected, the correlations between the various FAMALL versions are very high, with coefficients of 0.996 and higher using the complete sets of matrices for all overlapping years combined. The migration matrices from FAMALL sources do not correlate as well with those from INCTAX sources. The correlation coefficients involving the FAMALL migration matrices and the various types of INCTAX data, are as follows:

- with old UI data base for 1968-70	0.865
- with Winer and Gauthier data for 1968-77	0.960
- with new UI data base for 1972-82	0.751
- with Stats. Can. new data (91-210) for 1976-85	0.982.

These results do not present any major surprises, except that the degree of correlation between the two sets of migration matrices produced by Statistics Canada is high. This high correlation is probably related to the fact that Statistics Canada has used a set of parameters derived from the INCTAX source to estimate place-to-place flows for total population based on the FAMALL data (Statistics Canada 1987).

In view of the different patterns of the migration rates derived from CENS and FAMALL data (see Figure 1) we did not expect a high level of correlation between the matrices, but the actual coefficient, at 0.959, is surprisingly high. For purposes of this correlation analysis the annual FAMALL matrices for the relevant 5-year periods have simply been summed, to allow a comparison with the CENS data. Our surprise at the high correlation coefficient in part arises from the different treatment of repeat migrants in the CENS data based on 5-year time spans.

Migration matrices may be well correlated but still result in quite different net flows. Table 1 attempts to summarize the character of matrices in terms of this gross-net distinction. The ratio, R1, is computed for each matrix, where R1 is the sum of the net migration elements (positive numbers only) divided by the sum of the gross flow elements (Vanderkamp 1977). R1 has a 0-value when all pairwise migration flows are equal (all net migration elements are 0) and it equals 1 when all gross flows are uni-directional (half of the gross migration elements are 0). Table 1 reports the R1 ratios for the years 1961-2 to 1985-6 for the various migration sources; the data from the 1986 Census will not be available until 1989.

It may come as a surprise that net migration is not infrequently only 10 to 20 percent of gross migration. The R1 ratio tends to be larger when repeat migration has been largely "netted out", as in the case of the CENS figures. The ratios for FAMALL data are always smallest, with none of the numbers exceeding 25 percent. INCTAX data (1), (2) and (4) relate to the subgroup of tax filers in the total

population. Particularly interesting is the fact that the ratios for INCTAX (3) (STATCAN), which are adjusted to refer to total population, are always larger than those for FAMALL. We have already seen that these two series have similar migration rates and are fairly highly correlated. This suggests that repeat migrants (presumably within a 12-month period) are a fairly important component of FAMALL statistics. The R1 ratios display a kind of cyclical pattern, although not the typical business cycle, and they appear to be highest in the 1976-81 period, when the energy boom in the West resulted in large redistributions of population.

Table 1

RATIOS OF NET VERSUS GROSS MIGRATION FLOWS (R1)*

	FAMALL (STATCAN)	INCTAX** (1)	(2)	(3)	(4)	CENS
1961-62	.08					
1962-63	.11					
1963-64	.12					
1964-65	.15					
1965-66	.18					
1966-67	.16	.34				
1967-68	.12					
1968-69	.15		.18			.33
1969-70	.24	.30	.28			
1970-71	.22	.25	.23			
1971-72	.17	.33	.19			
1972-73	.17		.19		.38	
1973-74	.15		.20		.31	.27
1974-75	.14		.18		.25	
1975-76	.15		.22		.37	
1976-77	.15		.20	.22	.32	
1977-78	.19		.24	.26	.28	
1978-79	.20			.25	.32	.40
1979-80	.25			.32	.42	
1980-81	.23			.30	.30	
1981-82	.14			.20	.25	
1982-83	.10			.15	.36	
1983-84	.19			.22		
1984-85	.14			.17		
1985-86	.14			.17		

* Ratio R1 is sum of positive net migration elements divided by sum of migration matrix.

** The four INCTAX series are keyed as follows: (1) Old UI data base; (2) Winer & Gauthier annual tabulations, relate to the first calendar year, i.e. first listed ratio (0.18) refers to 1968; (3) Statistics Canada (#91-210) published data; (4) New UI data base.

Impacts of Migration

This discussion about net-gross ratios leads quite naturally to an evaluation of the impacts of migration in terms of population redistribution. Gross migration flows may be large without having much impact on the geographical distribution of population, and the discussion about these ratios has already indicated that the net flows are small by comparison with the gross flows.[7] We are by no means implying that small ratios and small impact measures indicate that migration is somehow ineffective. Some people argue that the prevalence of migration flows in opposite directions is an indication that migration is not an efficient process in relocating a country's labour resources, but this is clearly an untenable position. Migration decisions are motivated by a wide variety of factors, and individuals have different sets of tastes, talents and opportunities. The fact that individuals are moving in opposite directions may make perfectly good sense for each one concerned. Even repeat and return migration, which affect the net-gross impact measures, may well be "rational" given uncertainties and the costs of information. Nevertheless, impact measures of migration are quite legitimate as a way of evaluating different migration statistics.

So far migration data derived from the so-called survival method (SURV) have not featured in our comparisons, since the SURV method does not generate migration matrices but only net changes due to migration. Our comparisons in this section concern a broad measure of the impact of migration, the so-called net impact rates, and three case studies that deal with the impact of migration on three regions in the recent past. The migration data obtained from the SURV method can obviously be used in these comparisons. In fact it can be argued that in an *ex post* fashion the SURV data constitute the most comprehensive measure of the impacts of migration, provided we have confidence in the accuracy of the (census) measures of population and the data on births, deaths, immigration, and emigration that constitute its basic ingredients.[8]

[7]Typically migration models implicitly allow gross flows to exceed net flows. The traditional gravity model of migration in fact restricts gross flows in opposite directions to be the same, thus resulting in zero net flows and changes (Vanderkamp 1977).

[8]This point is related to the so-called error of closure test, which is widely used in *Population Estimation* (Statistics Canada 1987); the basic idea is that one should choose the estimation method, and data source, which minimizes the gap between estimated and actual population figures. In the context of our data series this implies that our SURV(DOM) series—estimated net internal migration data obtained by adjustment for net international migration—should be the standard against which other series are to be measured, since the SURV-data per definition are based on zero errors of closure. As *Population Estimation* makes clear, there are serious problems in applying this technique, basically because one is testing for the accuracy of a number of provincial population change components at the same time. The census counts of population are hampered by varying degrees of undercoverage, and the provincial estimates of immigration and emigration are of questionable quality.

The only problem is that the SURV statistics contain also the effects of net immigration or emigration. To modify our SURV statistics for the purpose of the following comparisons, we therefore use a set of Statistics Canada estimates of net immigration (or emigration) by province, estimated on the basis of INCTAX sources. These net immigration data are available back to 1961, and they were supplied by Statistics Canada on request. In the following analysis the resulting data series is labelled SURV(DOM) to indicate that the estimated provincial population changes are due to net domestic migration. The SURV series, which includes the net effects of domestic *and* international migration, is also shown, and a comparison of the SURV and SURV(DOM) series indicates the extent to which net immigration is important in the population distribution process. For the period 1981 to 1986 the SURV data are based on the post-censal estimates of population by province, since the inter-censal estimates of population for this period were not available at the time.

Figure 2 shows net impact rates for six time series of migration statistics, all of them clearly labelled at the bottom of the figure. The net impact rate measures net interprovincial migration (positive net figures only, with 10 provinces) as a percentage of the relevant Canadian population; for the SURV series the net numbers refer to net *total* migration, including net immigration. All six series in Figure 2 relate to total population, which should make them quite comparable, although the time dimension of the CENS series is, of course, different. The net impact rates for the five annual data series are in the 0.0 to 0.5 percent range, without much evidence of any kind of trend. The FAMALL(W&G) series links quite well with the INCTAX(STATCAN) series. The FAMALL and INCTAX series appear to be strongly correlated with the SURV(DOM) series, especially for the period since 1966.

The CENS data, of course, cover 5-year periods and, not surprisingly, the rates are substantially higher than the annual rates. The net impact rates from the CENS source display a pattern that is not unsimilar to those of the annual data, although a formal correlation (with only five data points) would not be strong. There appears to be a cyclical pattern in the annual impact rates and, while the pattern bears some resemblance to a business cycle, there are some departures from this, particularly in the last 15 years. In general the time patterns of the annual net impact rates, in Figure 2, are similar to the time patterns

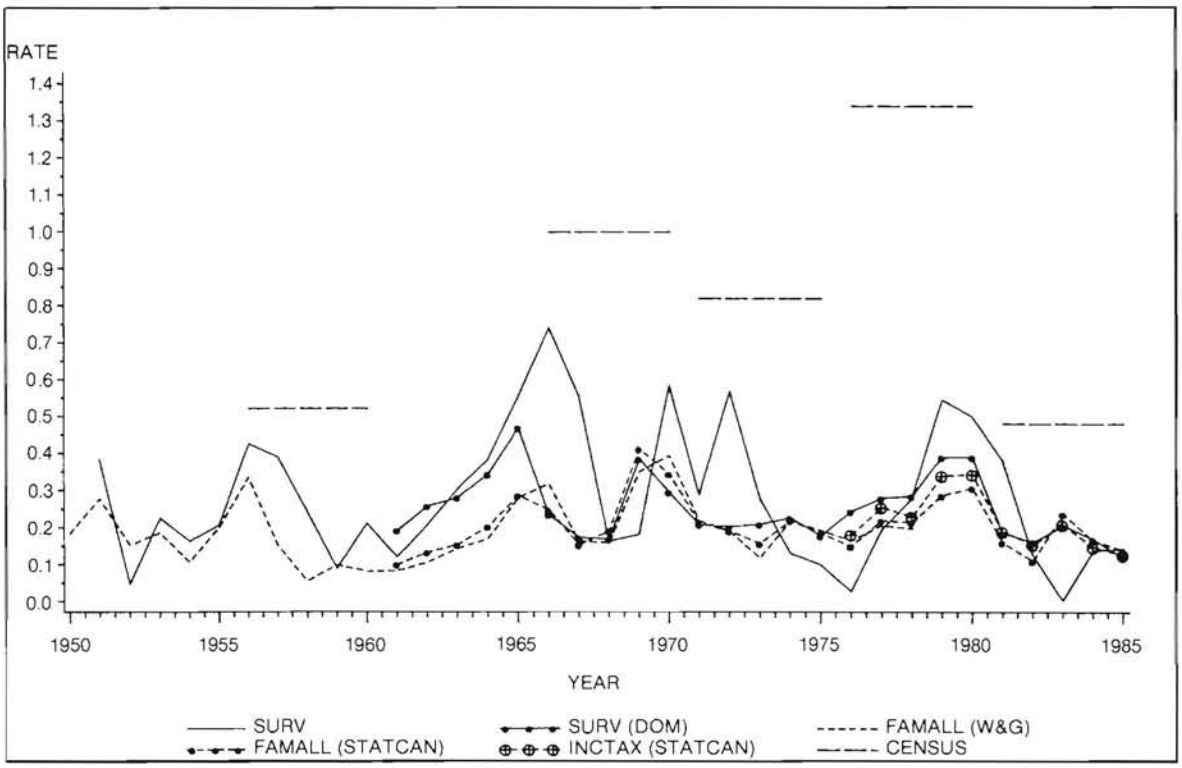

Figure 2

NET IMPACT RATES
(PERCENTAGE OF BASE YEAR POPULATION)

displayed by gross migration rates in Figure 1. There may be some surprise about the absolute size of the net impact rates, but it should be remembered that even with a 0.2 percent impact rate 2 percent of the population may be relocated after a decade.

The SURV series in Figure 2 is, of course, different from the others, as it includes the net impact of international migration as well. It is interesting to note that the time pattern of these SURV data shows considerably more variance than those of the other annual series. This suggests that net immigration is a major force in the process of population redistribution in Canada, and we shall see some specific examples of this in the case studies below.

Table 2 deals with our first case study, which raises the question of how the different data sources measured the alleged decline in net outmigration from the Atlantic provinces during the 1970s. A number of explanations have been advanced to explain this decline, from a general trend towards de-urbanization to the effects of specific policy initiatives such as UI reform and the system of DREE grants.

The data series in Table 2 are agreed that there was some change in the migration pattern involving the Atlantic region in the early seventies, but there is considerable disagreement about the size of the shift, with estimates ranging from 10,000 to 90,000 as the change in net flows between 1966-71 and 1971-76. Moreover, the CENS and SURV(DOM) data indicate that this change in pattern was fully reversed in the 1976-81 period, but the FAMALL (STATCAN) series shows much less of turnaround.

The migration pattern for the most recent 5-year period, 1981-86, is rather like that for the early seventies, except that the FAMALL (STATCAN) data show net outmigration from the Atlantic provinces compared with substantial net inmigration in the 1971-76 period. There is rather remarkable agreement between the INCTAX (STATCAN), CENS, and SURV(DOM) data about the patterns of the last decade. The SURV data, which include net immigration, indicate that international migration has reinforced the tendencies of domestic migration patterns. It could be that in both these periods somewhat depressed general economic conditions have had more to do with the Atlantic migration pattern than any specific policy developments.

Table 3 deals with the question of how the different sets of migration statistics have reflected the major changes in migration patterns involving Alberta in the last fifteen years. The boom-bust turnaround in the energy sector has had a major impact on economic opportunities in Alberta and therefore presumably on migration.

There is substantial agreement among the series presented in Table 3 that the major increase in net migration into Alberta commenced in 1974-75. There is also rough agreement that net migration

Table 2

NET INMIGRATION INTO ATLANTIC PROVINCES 1966-86
('000 of persons)

Data Sources	NFLD	PEI	NS	NB	TOTAL*
FAMALL (STATCAN)					
1966-71	-19	-3	-16	-20	-58
1971-76	- 2	4	11	17	30
1976-81	- 8	1	0	4	- 3
1981-86	-15	0	5	- 2	-12
INCTAX (STATCAN)					
1976-81	-19	-1	- 7	-10	-37
1981-86	-15	1	7	0	- 7
CENS					
1966-71	-15	-1	- 8	- 8	-32
1971-76	- 7	2	5	9	10
1976-81	-20	0	- 8	- 9	-37
1981-86	-17	2	6	- 1	-10
SURV (DOM)					
1966-71	-11	-1	3	- 3	-12
1971-76	- 9	1	2	4	2
1976-81	-26	0	- 7	- 8	-41
1981-86	-14	1	7	- 1	- 7
SURV (including net immigration)					
1966-71	-20	-2	- 4	-17	-43
1971-76	- 5	2	7	10	14
1976-81	-25	0	- 9	-10	-44
1981-86	-11	1	15	3	8

* Totals may not add due to rounding.

during the period 1976-81 was at a high level, but there is disagreement about the amount, with estimates ranging from about 150,000 to 240,000, certainly an extraordinarily high level by any count. The higher numbers for the SURV series indicate that net immigration to Alberta was about 35,000 in the same 5-year period, which does not seem a very high level. There is not full agreement between the INCTAX(NEW UI) and other data about the timing of the turn-around in the eighties. According to the INCTAX(NEW UI) source, the decline in the inmigration rate had already started in 1981-82 and the rate of net outmigration from Alberta was already in the 30,000 range in 1982-83, but the other series agree that the turn-around came in 1982-83 and that the large outmigration did not start until

Table 3

NET MIGRATION INTO ALBERTA 1971-86
('000 of persons)

	FAMALL (STATCAN)	CENS		INCTAX 2 Series NEW UI	INCTAX 2 Series STATCAN	SURV (DOM)	SURV
1971-72	4					7	10
1972-73	6			-15		9	14
1973-74	2			7		8	14
1974-75	23			22		27	37
1975-76	25			21		28	39
1971-76	59	62	'72-76	(35)		79	114
1976-77	25			20	35	46	53
1977-78	26			27	33	43	47
1978-79	30			38	33	44	46
1979-80	31			45	41	52	62
1980-81	38			39	44	55	69
1976-81	151	198		169	186	240	276
1981-82	41			7	36	37	49
1982-83	- 3			-30	-12	-12	1
1983-84	-43				-32	-32	-34
1984-5	-27				-21	-27	-32
1985-6	- 1				- 4	- 8	- 3
1981-6	-33	-28			-33	-42	-20

1983-84, reaching a reduced net outflow of only a few thousand in 1985-86.

The last case deals with the migration experience of British Columbia during the same period. BC has always had large swings in net inmigration, with years of net outflow suceeding periods of large net inflows. This up-and-down pattern is associated with general economic conditions in Canada and in BC in particular. Table 4 certainly provides plenty of evidence of this roller coaster pattern of BC migration.

When the energy boom began, the impact on BC migration went in the opposite direction to that in Alberta. Between 1973-74 and 1975-76 net migration into BC decreased by about 40,000, and the various series are in agreement on that score. In 1979-80 the net rate of inmigration into BC peaked, with an increase of about 40,000, although the INCTAX(NEW UI) data make the shift somewhat sharper, with more annual variation as well. During Alberta's high growth period from 1976 to 1981, BC also had a large amount of net inmigration of more than 100,000 people, and the various sources produce rough agreement on this number; but this experience was not a sharp departure for BC, as the rates of net inmigration in 1976-81

Table 4

NET MIGRATION INTO BRITISH COLUMBIA 1971-86

	FAMALL (STATCAN)	CENS		INCTAX 2 Series NEW UI	INCTAX 2 Series STATCAN	SURV (DOM)	SURV
1971-72	27					31	40
1972-73	27			26		31	45
1973-74	30			26		38	57
1974-75	12			-21		18	41
1975-76	-4			-18		1	16
1971-76	92	96	'72-76	(13)		119	199
1976-77	2			13	5	6	15
1977-78	16			3	18	19	25
1978-79	20			16	22	23	28
1979-80	39			38	40	41	57
1980-81	38			27	38	39	57
1976-81	115	111		97	123	128	182
1981-82	7			- 8	9	9	24
1982-83	4			- 1	-1	-1	11
1983-84	13				7	7	24
1984-85	-2				-2	-2	-1
1985-86	-8				-5	-8	16
1981-86	14	10			8	5	73

were not very different from those of 1971-76. At the same time, net international migration may have contributed in the order of 60,000 to 80,000 persons to BC's population. While BC did not share the sharp turn-around of Alberta during the early 1980s, inmigration into BC was at a much lower level with considerable year-to-year variation, a pattern on which the various available data series appear to agree, broadly speaking.

Conclusion

Migration researchers in the U.S. have expressed envy about the availability of Canadian migration statistics, particularly in relation to our long time series of provincial migration matrices, but as we have seen, the FAMALL data have some quite serious deficiencies, and Statistics Canada has reduced their prominence in its own work and has stopped publishing them. Statistics Canada's final post-censal estimates of provincial populations are now based on data derived from INCTAX sources, and these appear to be more reliable, but consistent series are only available for the last ten years. A micro data base

derived from INCTAX sources (with T1 form information) would be very useful, but no such data base is currently available.

We have no reason to believe that CENS migration data are unreliable, but we are left with a few minor puzzles about the behaviour of gross migration and net impact rates for CENS data compared with those of other series. The CENS micro data sets are probably the best currently available in Canada, but their usefulness is somewhat limited as this is not a longitudinal data base.

The SURV data set should probably only be employed when other data are not available and when the analytical goal is to work with a comprehensive supply change concept.

The lack of a good longitudinal micro data base is the most serious problem in Canadian migration research. We have nothing comparable to the major U.S. micro data sources. We therefore conclude with a plea for the establishment of a comprehensive panel survey with socio-economic data of all kinds, including detailed data on migration.

References

Economic Council of Canada. 1977. *Living Together*. Ottawa.

Grant, E. Kenneth, and John Vanderkamp. 1986. "Repeat migration and disappointment", *Canadian Journal of Regional Science*, 9:299-322.

Grant, E. Kenneth, and John Vanderkamp. 1985. "Migrant information and the remigration decision: Further evidence", *Southern Economic Journal*, 51:1202-1215.

Grant, E. Kenneth, and John Vanderkamp. 1984. "A descriptive analysis of the incidence and nature of repeat migration within Canada, 1968-71", *Canadian Studies in Population*, 11:61-78.

Grant, E. Kenneth, and John Vanderkamp. 1983. "Regional demand-supply projections and migration", in *Canadian Labour Markets in the 1980's*. Proceedings of a Conference held at Queen's University. Kingston: Industrial Relations Centre, Queen's University, 112-127.

Grant, E. Kenneth, and John Vanderkamp. 1980. "The effects of migration on income: A micro study with Canadian data, 1965-71", *Canadian Journal of Economics*, 13:381-406.

Grant, E. Kenneth, and John Vanderkamp. 1976. *The Economic Causes and Effects of Migration: Canada. 1965-71*. Ottawa: Economic Council of Canada.

Greenway, H. F., and G. W. Wheatley. 1963. "Regional aspects of labour mobility in Canada, 1956-59", in *Papers of the 1961 C.P.S.A. Conference on Statistics*, edited by Wm. C. Hood and John A. Sawyer. Toronto: University of Toronto Press.

Kasahara, Yoshiko. 1963. "The flow of migration among the provinces in Canada, 1951-61", in *Papers . . .* (see Greenway and Wheatley 1963.)

Liaw, Kao-Lee, and Pavlos Kanaroglou. 1983. "Metropolitan elderly outmigration in Canada, 1971-76", in *Research on Aging*, 8.

Liaw, Kao-Lee, and Jacques Ledent. 1988. "Joint effects of personal and ecological factors on elderly interprovincial migration in Canada", *Canadian Journal of Regional Science*, 11:79-102.

Marr, W. L., and F. W. Millerd. 1988. "Migration and the employment status of married women", *Canadian Journal of Regional Science*, 11:121-134.

Marr, W. L., and F. W. Millerd. 1987. "Migration and changes in employment". Wilfrid Laurier University working paper.

Marr, W. L., and F. W. Millerd. 1980. "Employment income levels of interprovincial migrants versus non-migrants, Canada, 1971", in J. L. Simon and J. DaVanzo (eds.), *Research in Population Economics*, Vol. 2. Greenwich, Conn.: J.A.I. Press.

Nickson, May. 1967. *Geographic Mobility in Canada, October 1964-5*. Statistics Canada, Special Labour Force Studies #4.

Robinson, Chris, and Nigel Tomes. 1982. "Self-selection and interprovincial migration in Canada", *Canadian Journal of Economics*, 15:474-502.

Shaw, R. Paul. 1985. *Intermetropolitan Migration in Canada*. Toronto: N.C. Press.

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

Statistics Canada. 1983. "On the move: Results of a special survey on migration", in *The Labour Force* (71-001).

Statistics Canada. 1982. "Characteristics of migrants to Alberta and British Columbia, 1976-80", in *The Labour Force* (71-001).

Vanderkamp, John. 1988a. "Regional disparities: A model with some econometric results for Canada", in Benjamin Higgins and Donald J. Savoie (eds.), *Regional Economic Development: Essays in Honour of François Perroux*. Boston: Allen and Unwin, 269-296.

Vanderkamp, John. 1988b, forthcoming. "The role of migration in regional adjustment," in H. W. Herzog and A. M. Schlottmann (eds.), *Migration and Labour Market Efficiency*. The Netherlands: Martinus Nijhoff.

Vanderkamp, John. 1984. "The efficiency of the interregional adjustment process", in Kenneth Norrie (research coordinator), *Disparities and Interregional Adjustment*, Vol. 64 of the research studies prepared for the Royal Commission on the Economic Union and Development Prospects for Canada. Toronto: University of Toronto Press, 53-108.

Vanderkamp, John. 1977. "The gravity model and migration behaviour: An economic interpretation", *Journal of Economic Studies*, 4:89-102.

Vanderkamp, John. 1973. *Mobility Behaviour in the Canadian Labour Force*. Special Study No. 16. Ottawa: Economic Council of Canada.

Vanderkamp, John. 1971. "Migration flows, their deteminants and the effects of return migration", *Journal of Political Economy*, 79:1012-1031.

Vanderkamp, John. 1968. "Interregional mobility in Canada: A study of the time pattern of migration", *Canadian Journal of Economics*, 1:595-608.

Winer, Stanley L., and Denis Gauthier. 1982. *Internal Migration and Fiscal Structure*. Ottawa: Economic Council of Canada, Supply and Services Canada.