Canadian Studies in Population, Vol. 6, 1979, pp. 33-44

TRACING RESIDENTIAL MOBILITY THROUGH DATA LINKAGE OF TAX ASSESSMENT RECORDS: AN EXAMPLE OF RESIDENTIAL STABILITY OF ALIENS IN A MID-ONTARIO METROPOLITAN AREA

Mary Thompson, Daniel Kubat, Frank Fasick, Steven Hawkins
University of Waterloo, Waterloo, Ontario, Canada

Résumé — Les rôles des contributions en Ontario sont à présent comparés et traités par l'ordinateur. Comme tels, ils représentent une base des données (et une base de sondage) qui offre de nouvelles occasions pour l'analyse de la mobilité intraurbaine et intraprovinciale . On a essayé de démontrer l'utilité de ces données en présentant des conclusions explicatives sur la mobilité des ménages ayant un étranger comme chef, à l'intérieur des cités jumelles de Kitchener-Waterloo entre 1972 et 1973. De plus, on a noté l'utilité des données comme aide au développement et à l'essai des modèles de provenance mathématique pour la prédiction des taux de mobilité. On a décrit un exemple d'un tel modèle et on a présenté un éclaircissement en se fondant sur les données pour la période 1972-1973, données employées dans cette étude.

Abstract —Tax assessment rolls in Ontario are now standardized and processed by computer. As such they represent a data base (and sampling frame) that opens up new opportunities for the analysis of intra-urban and intra-provincial mobility. An attempt is made to demonstrate the utility of this data by presenting illustrative findings on the mobility of households with an alien head within the twin cities of Kitchener-Waterloo between 1972 and 1973. In addition, the usefulness of the data in aiding the development and testing of mathematically derived models for predicting mobility rates is noted. An example of such a model is described and an illustration of its application is presented making use of the 1972-73 data utilized in this paper.

Key Words — **residential mobility, linkage, assessment records, aliens**

This paper[1] is a report by the Kitchener-Waterloo Metropolitan Area Study Group on the usefulness of tax assessment rolls for the study of intra-city and intra-provincial migration within Ontario. In addition to briefly describing the tax assessment records, we present an illustrative analysis based on the movement of aliens within Kitchener-Waterloo between the years 1972-73. So far as we know, this is the first time such data has been reported for aliens in Canada.[2] The analysis includes the description of a simplified version of a more complex model of migration currently being used in a five-year study of migration within Kitchener-Waterloo being carried out under the direction of one of the authors.

Over the last several years there has been a rapid growth of published studies on intra-urban migration. A major part of the work has been done by geographers or econometricians, although there has been increasing interest by sociologists. The latter have been interested primarily in the whys of the moves (Rossi, 1955); the circumstances of the moves — such as stage in the life cycle (Leslie and Richardson, 1968; Maisel, 1966); residential preferences (Michelson et al., 1972); and the characteristics of the movers (Stone, 1970, 1971). Both sociologists and geographers, however, have recently paid greater attention to describing intra-urban moves in terms of mathematical models. Important work in this area has been done by Simmons (1974, 1973, 1968), as well as by Myers et al. (1965), Brown and Moore (1970), and others.

Mary Thompson, Daniel Kubat, Frank Fasick and Steven Hawkins

I. *The Provincial Tax Assessment Records*

Historically an important stumbling block to the study of intra-urban migration in North America has been the lack of adequate data. We have found potentially very useful data for the study of intra-urban and intra-provincial migration in the province of Ontario tax assessment records. They are the foundation stone of our efforts to study migration within Kitchener-Waterloo.

Ontario tax assessment records provide a complete listing of households and of the residents in the households, whatever their income and whether or not they pay property tax. In recent years the records have been standardized and stored on computer tape. Consequently, they provide a readily retrievable (and editable) listing of the population and of households within the province.

The usefulness of the tax assessment records is enhanced by the fact that a number of important demographic characteristics of individuals appear in the records. These include sex, year of birth, and whether or not the individual is the household head. Additional information includes citizenship, and if household head, whether owner or tenant. It is also possible to construct from the records information concerning households, such as size, composition, and type of housing.

The appearance of these variables on the records means that it is feasible to sample efficiently a number of important demographic sub-groups within the population of Ontario for survey research studies.

Because the tax rolls are continuously updated in terms of housing stock, etc., and updated yearly in terms of residents, they are particularly suitable for the study of migration over time.

The amount of error in the tax assessment records is an important consideration when contemplating the use of the records for research purposes. We have not done a thorough study of the error in the records, but we have done some preliminary analyses. We conducted enumerations of households in several areas and discovered that the tax listing was more complete than the one obtained by our enumerators. We have also done checks based on inconsistencies in the records. These mostly involve changes in spelling of names and addresses from one year to the next, inconsistencies which are important when attempting to determine the movement of persons and households from year to year through computer matching. Based on our preliminary analyses, inconsistencies from one year to the next occur in something like five per cent of the records. In the data on aliens presented below, we have utilized a correction for this type of error.[3] In general these errors cause mobility to be overestimated, since an apparent address change is taken to indicate a move within the community, while a change in name is classified by the matching program as a departure from the community.

An additional source of error occurs when residents who move fail to report that they have left their old residence when they return their tax assessment form from their new residence. These residents are enumerated twice (or more) in one year, once for each address, and the move may not be counted for that year. This source of error tends to pull mobility estimates downward, and has not been taken into account in our example. An analysis of this source of error has shown that it affects about three per cent of cases, and that it can largely be eliminated through improvements in the matching program.

II. *Illustrative Analysis of Residential Mobility among Aliens in Kitchener-Waterloo*

To illustrate the usefulness of the tax records in studying intra-urban (or alternatively, intra-provincial) residential mobility, we are presenting a brief analysis of alien movers and non-movers within Kitchener-Waterloo between 1972 and 1973. This analysis of migration among aliens grew out of a study conducted by Daniel Kubat for the Office of the Secretary of State, Ottawa.

There are two parts to the analysis we offer below. First, we present some preliminary descriptive findings regarding mobile and non-mobile households with alien heads in Kitchener-Waterloo between 1972 and 1973. Second, we present a brief discussion of a model that makes use of the time series characteristic of the tax assessment records in predicting mobility rates and use the 1972-73 data on aliens to provide an example.

There were 3,257 households with an alien head in the 1972 tax rolls. Table 1 shows the residential mobility of these households between 1972 and 1973 for both Kitchener and Waterloo. In both communities the mobility rate was quite high with nearly 50 per cent of the households engaged in some sort of move; more specifically, about 35 per cent moved outside Kitchener-Waterloo and about 15 per cent moved within the cities. In 1973 about 42 per cent (1,545/3,705) of alien households were newcomers to the list.

TABLE 1 RESIDENTIAL MOBILITY BETWEEN 1972-1973 OF HOUSEHOLDS WITH ALIEN HEADS KITCHENER-WATERLOO

Mobility Status	Kitchener	Waterloo	Total
Did not move	52% (1287)	51% (400)	52% (1687)
Moved within K-W	15% (361)	14% (112)	15% (473)
Moved outside K-W	33% (828)	35% (269)	33% (1097)
Total present in 1972	2476	781	3257
Moved into K-W 1973	1062	483	1545
Total present in 1973	2685	1020	3705

Table 2 shows the turnover in occupancy status (whether tenant or owner) among households with an alien head moving within Kitchener-Waterloo between 1972 and 1973. Mobility is clearly associated with assuming home ownership. Only a little over one-quarter of the mobile households owned their home in 1972, but more than one-half did so in 1973.

Altogether, 33 per cent of the mobile households changed their status from tenant to owner between 1972 and 1973. This figure is nearly as large as the 40 per cent of households moving as tenants.

These figures fit with the frequent observation that immigrants place home ownership high on their list of priorities.

Table 3 shows a marked and consistent pattern of reduced residential mobility with increasing age of head among alien households. This pattern is primarily the result of a

TABLE 2 TURNOVER IN OCCUPANCY STATUS FROM 1972-1973 AMONG RESIDENTIALLY MOBILE HOUSEHOLDS WITH AN ALIEN HEAD KITCHENER-WATERLOO

	Occupancy Status of Households Moving Within Kitchener-Waterloo Between 1972 and 1973		
1972	1973		Totals
	Tenant	Owner	1972
Tenant	40%	33%	73%
Owner	05%	22%	27%
Totals - 1973	45%	55%	100% (458)

Note: 15 cases for which complete data were not available are excluded.

TABLE 3 RESIDENTIAL MOBILITY BETWEEN 1972-1973 OF HOUSEHOLDS WITH ALIEN HEADS BY AGE OF HOUSEHOLD HEAD KITCHENER-WATERLOO

Mobility Status	Age of Household Head					
	11-19	20-29	30-39	40-49	50-69	70+
Did not move	--	36%(219)	51%(675)	61%(177)	71%(358)	77%(86)
Moved within K-W	--	18%(110)	15%(207)	13%(37)	10%(49)	5%(5)
Moved outside K-W	100%(3)	46%(281)	34%(451)	26%(77)	19%(97)	18%(20)
Total present in 1972	3	610	1333	291	504	111
Moved into K-W, 1973	17	609	544	78	103	25
Total present in 1973	17	938	1426	292	510	116

Note: 74 cases for which complete data were not available are excluded.

marked decrease in households moving outside (and into) Kitchener-Waterloo with increasing age of head. Movements within Kitchener-Waterloo remain relatively stable as the age of the household head increases.

This table indicates the importance of in- and out-migration on patterns of residential mobility by age among households with an alien head.

TABLE 4 PERCENTAGE OF MOVERS WITHIN KITCHENER-WATERLOO BETWEEN 1972-1973 AMONG HOUSEHOLDS WITH ALIEN HEADS BY CITIZENSHIP STATUS AND AGE

	Percentage of Households Moving within Kitchener-Waterloo Between 1972 and 1973 among Households in Which the Head Was an Alien in 1972 and in 1973 was:	
Age	Still an Alien	A Citizen
20-29	27%	58%
30-39	21%	39%
40-49	16%	22%
50-69	10%	19%
70+	05%	67%

Note: 64 cases for which complete data were not available are excluded.

In Table 4, we relate residential mobility within Kitchener-Waterloo to change in citizenship of the household head between 1972 and 1973, and his age. Within the 20 to 29 age group, well over one-half of the households in which the head changed his citizenship moved compared with little more than one-quarter of households in which the head remained an alien. Among the 30 to 39 age group the relationship is in the same direction but somewhat reduced. For the older age groups the figures are in the same direction but the differences are negligible. The floor effect probably has some effect on the small differences within the older age groups, but it is clear that the relationship between change in citizenship of household head and intra-city mobility is specified by age.

Younger heads of households are likely to be undergoing relatively frequent changes in status. Often these changes are related to residential mobility, such as marriage or switching jobs. It appears that acquiring citizenship within these age groups is tied in with such patterns.

We hope that the preliminary findings we have presented are sufficient to demonstrate the usefulness of the tax assessment records in studying geographic mobility. Not only do they provide some important basic information directly as we have shown, but they provide an efficient sampling frame for migration studies based on survey research. Such studies are needed to obtain factual and attitudinal data in the depth necessary to develop adequate theories of short-range geographic mobility.

III. Modelling Intra-urban Mobility

Geographic mobility is one of the areas in which social scientists have made serious attempts to move beyond verbal theories by expressing propositions in mathematical terms. As part of our population study, we have been investigating the implications of modelling yearly changes in type or location of residence as a discrete semi-Markov process.

It has been borne out by empirical studies of intra-urban migration (see, e.g., Simmons, 1974) that simple Markov chains are inadequate for modelling transitions of families among residence types. One of the main reasons is that the probabilities for timing and destination of a move appear to depend on how long the family has been in its current residence. Indeed, several mobility studies (e.g., Morrison, 1967; Land, 1969) have supported the axiom of cumulative inertia, that the probability of moving tends to decrease with length of stay. Whether the cumulative inertia axiom is applicable or not it seems reasonable to expect a dependence of mobility probabilities on duration of sojourn. The idea of using semi-Markov models, in which this sort of dependence can be incorporated, seems to have appeared fairly recently, for example, in papers by McGinnis (1968), Ginsberg (1971), and Gilbert (1972).

The first step in specifying a semi-Markov model is to determine a state space. For example, if movement of families among three geographical zones is of interest, this state space could consist of the following four states:

$Z1$ = *living in Zone 1*

$Z2$ = *living in Zone 2*

$Z3$ = *living in Zone 3*

∇ = *dissolution of family or leaving the area for which data are available.*

Alternative state spaces may be formed by subdividing and/or combining states.

Once the state has been determined, we denote by p_{ij} the probability that a family now entering state i will make its next move to state j. Clearly the sum j of over p_{ij} is 1. In a semi-Markov model it is assumed that this probability does not depend on the (residential) history of the family before it enters state i. It is also assumed that the distribution of length of stay in state i, given that the next move is to state j, does not depend on the history of the family before it enters state i. We may denote by $\phi_{ij}(x)$ the probability that this length of stay is x years, $x = 1, 2, \ldots$. In a Markov chain model, any such distribution would be geometric, with $\phi_{ij}(x)$ having the form $(1-\theta_i)^{x-1}\theta_i$; in the semi-Markov case, the frequency function ϕ_{ij} can have any form.

For example, suppose the discrete semi-Markov model with the above state space is considered to be applicable, and suppose that it is required to compute the probability that a family entering the population in $Z1$ in 1960 moves to another residence in Zone 1 in 1962, moves to one in Zone 3 in 1963, and then leaves the population in 1966. In terms of p_{ij} and ϕ_{ij} this probability will be given by the product

$$p_{Z1Z1}\phi_{Z1Z1}(2)p_{Z1Z3}\phi_{Z1Z3}(1)p_{Z3\nabla}\phi_{Z3\nabla}(3)$$

of six factors, with each successive pair of factors relating to one transition, and being determined independently of preceding pairs.

Varying the form of the frequency function $\phi_{ij}(x)$ may make the axiom of cumulative inertia or its opposite (or neither) obtain. It will obtain for a given transition type (i to j) if the conditional probability of moving within the next year, given no move up to year n after entry into i, decreases as a function of n. In symbols, cumulative inertia will hold (for an $i-j$ transition) if

$$\phi_{ij}(n+1)/[1-\sum_{x=1}^{n}\phi_{ij}(x)]$$

is a decreasing function of n.

A further assumption needed to simplify the statistical analysis at the estimation stage is that the inflow of families from outside the area of the data into each state is stationary over time. For practical purposes, the stationary quality of inflow over the recent past is what is required.

Even with just two years of data, it is possible using these assumptions to make some useful interpretations. The observed data include (a) the number (a_i, say) of families in state i in Year 1, for each state i (but no information on their sojourn time to date in their present residences); and (b) the number (b_{ij} say) of the a_i families in state i in Year 1 who move to state j by Year 2. If (for mathematical convenience) stationary inflow over all past years is assumed, with N_k being the expected number of families entering the population in state k in each year, the expected value of a_i is $Ne_i\mu_i$ where

$N = \sum_k N_k$ = the expected number of families entering the population each year,

e_i = the expected number of transitions made by a randomly selected family into state i, and

$\mu_i = \sum_{x=1}^{\infty} x \sum_j p_{ij}\phi_{ij}(x)$ = the expected length of stay in state i, unconditional on the subsequent state j.

The fact that the expected number of families found in state i at a given time is proportional to the average number of stays (e_i) times the expected length of each such stay (μ_i) is intuitively clear for a steady-state population.

Similarly, the expected value of b_{ij} is Ne_ip_{ij}. Heuristically, this formula may be derived from the previous one by saying that of the families in state i in Year 1, about one in every μ_i is at the end of its stay in its present residence; of these imminent movers, a proportion, p_{ij}, may be expected to move to j. Thus b_{ij}/a_i, the proportion of families in state i in Year 1 which moves to state j by Year 2 estimates p_{ij}/μ_i. The rigorous derivation of these formulas is based on the elementary theory of semi-Markov processes, and some details are given in the appendix. It should be recalled that the transition probability p_{ij} does not relate to a one-year period or to any fixed period of time; it is the probability of *ultimate* transition to j, given entry into state i. Thus b_{ij}/a_i does not estimate p_{ij} itself, but a smaller quantity, namely p_{ij}/μ_i.

The estimate b_{ij}/a_i can also be interpreted as an estimate of the probability that a randomly selected family in state i will move to state j within a year. As a consequence, the proportion of families in state i in Year 1 who move anywhere by Year 2, measured by $\Sigma(b_{ij}/a_i)$, estimates $1/\mu_i$, or the reciprocal of the mean length of stay in years. Estimates of the p_{ij} (as $b_{ij}/\Sigma b_i$) can then be obtained.

It should be emphasized that the assumption of stationary inflow is introduced to make the problem of deriving estimates mathematically tractable. On a more intuitive level, the nature of this assumption can be seen by an examination of its bearing on the interpretation of the data. If the yearly residential histories were assumed to follow a homogeneous Markov chain, the quantity b_{ij}/a_i) would naturally estimate the corresponding one-step transition probability. However, as soon as the residential histories are seen as being non-Markovian, some assumption concerning the sojourn times of the a_i families in state i in Year 1 is required in order to interpret the b_{ij}/a_i as probabilities in terms of model parameters. The interpretation, in a semi-Markov context, of b_{ij}/a_i as an estimate of p_{ij}/μ_i, *or* as an estimate of the probability that a randomly selected state i family in any year will move to state j within a year, depends on the stationary inflow assumption.

Although the intuitive appeal of these interpretations justifies to some extent the adoption of the stationary inflow assumption, it should be remembered that in reality inflows, and indeed mobility within a community, are affected over the years by changes in economic and other environmental factors. Hence the assumptions of the semi-Markov model and the stationary quality of inflows can be expected to hold over periods of a few years at best. Estimates of model parameters must clearly be interpreted in this light. Nevertheless, Monte Carlo studies on artificial populations, to be reported in a forthcoming paper, indicate that the parameter estimates are robust to fairly substantial departures from the stationary inflow assumption.

With more than two years of data, it is possible also to estimate the distributions of lengths of stay in any state. Once these "parameters" of the model have been estimated, along with p_{ij}, it should be possible to drop the assumption of stationary inflows, and use actual past inflows to forecast future population development. A project incorporating this procedure is currently underway. Details of the estimation procedures and their possible use in forecasting housing demand will be given in a forthcoming paper.

To illustrate the mobility estimation procedure based on two years, we will use the Kitchener-Waterloo tax-roll data on alien movers and non-movers for 1972-73. We decided for purposes of illustration to make the state space consist of two "zones", Kitchener (K) and Waterloo (W), together with a "departure" state (∇), which would the-

TABLE 5 RESIDENTIAL MOBILITY BETWEEN KITCHENER AND WATERLOO FOR 1972-1973 BY HOUSEHOLDS WITH AN ALIEN HEAD

	Uncorrected	Corrected	Proportions
From Kitchener to Kitchener	302 households	284 households	.115
From Kitchener to Waterloo	59 households	59 households	.024
From Kitchener to Outside Twin Cities or dissolution	828 households	787 households	.317
Total Kitchener movers	1189 households	1130 households	.456
Total Kitchener non-movers	1287 households	1346 households	.544
Total Kitchener aliens (1972)	2476 households	2476 households	1.000
From Waterloo to Waterloo	78 households	72 households	.092
From Waterloo to Kitchener	34 households	34 households	.044
From Waterloo to Ontario Twin Cities or dissolution	269 households	256 households	.328
Total Waterloo movers	381 households	362 households	.464
Total Waterloo non-movers	400 households	419 households	.536
Total Waterloo aliens (1972)	781 households	781 households	1.000

oretically include household dissolution and emigration from the K-W area. A preliminary matching of the 1972 and 1973 rolls yielded the totals in the "uncorrected" column of Table 5.

It was noted that where the same address is written differently in 1972 and 1973, a non-moving household will be counted as a household moving within Kitchener or Waterloo. In cases where the household head's name is written differently in the two years, a non-moving or internally moving household will be counted as a departure and an arrival in 1973. It was estimated from a sample of the data that about five per cent of apparent internal movers are really non-movers, and that about five per cent of apparent departures are also non-movers. The "corrected" column of Table 5 deflates the proportions of movers accordingly.

In the last column of Table 5 are the proportions, computed separately for initial states Kitchener and Waterloo, corresponding to the corrected figures. For instance, the proportion of Kitchener alien households which moved within Kitchener between 1972 and 1973 was 284/2,476 or .115.

From the results in Table 5, it appears that if the semi-Markov model with stationary inflow were appropriate, the mean length of stay μ_k in Kitchener for an alien household would be 1/.456 or 2.19 years; the mean length of stay μ_w in Waterloo would be 1/.464 or 2.16 years. The estimated probability P_{kk} than an alien family now moving into Kitchener will make its next move (whenever that may be) within Kitchener would be .115 (which estimates P_{kk}/μ_k) times 2.19, or .252; the estimated probabilities for other destinations are given in Table 6, along with estimated standard errors.

TABLE 6 MOBILITY PARAMETERS p_{ij} FOR ALIEN HOUSEHOLDS

From:	To:	Kitchener		Waterloo		Departure
Kitchener	Pkk =	.252 ±.013	Pkw =	.052 ±.007	Pw∇ =	.696 ±.014
Waterloo	Pwk =	.094 ±.016	Pww =	.199 ±.021	Pw∇ =	.707 ±.025
	P∇k =	0	P∇w =	0	P∇∇ =	1

The sub-population of aliens clearly has very high geographic mobility, and because of this fact it is felt that the only assumption of our model which may be seriously in error is the assumption of stationary inflows. The numbers of arrivals in Kitchener and Waterloo in 1973 were 1,062 and 483 respectively, while the numbers of departures (as in Table 5) were 787 and 256. These figures suggest that the arrivals of aliens might have tended to increase in the early 1970s, and that the probabilities in Table 6 would be biased toward recent arrivals. That is, P_{kk} might better be interpreted as a probability that an alien household now moving to a Kitchener address and destined to move within a short time, will make its next move within Kitchener. This interpretation still carries with it valuable information about the movement of this sub-population within the area.

In this paper we have attempted to describe data that has recently become available and that opens up new opportunities for the analysis of intra-urban and intra-provincial mobility within Ontario.

The study of migration within Ontario making use of tax assessment records has only just begun. We are confident, however, that competent, careful, professional use of these records, with the co-operation of the Ontario government, can lead to meaningful advances in the study of intra- and inter-regional migration within Ontario. By inference and implication such work would hopefully contribute to our understanding of short-distance migration within North American and urban-industrial societies generally.

Footnotes

1 This article is a revised version of a paper presented at the Annual Meetings of the Canadian Population Society, University of Laval, Quebec City, May 26, 1976.

2 Because of tightened restrictions regarding access to computerized tax assessment records, this type of data is not currently available for academic research to the authors' knowledge. It is our hope that this report demonstrates the value of these data for responsible theoretical research and will facilitate renewed access to them.

3 The term "alien" is used on the assessment rolls to designate a person whose citizenship is other than Canadian.

References

Brown, L. A. and E. Moore. 1970. The intra-urban migration process: a perspective. Georgrafiska Annaler Series B, 52:1-13.

Gilbert, G. 1972. Two Markov models of neighbourhood housing turnover. Environment and Planning 4:133-146.

Ginberg, R. 1971. Semi-Markov processes and mobility. Journal of Mathematical Sociology 1:233-262.

Land, K. C. 1969. Duration of residence and prospective migration: further evidence. Demography 6:133-140.

Leslie, Gerald R. and Arthur H. Richardson. 1961. Life-cycle, career patterns, and the decision to move. American Sociological Review 26:894-902.

Maisel, S. J. 1966. Rates of ownership, mobility and purchase. In Essays in Urban Land Economics. Los Angeles: University of California Real Estate Research Program: 76-107.

McGinnis, R. 1968. A stochastic model of social mobility. American Sociological Review 33:712-721.

Michelson, William D. Belgue, and J. Stewart. 1972. Intentions and expectations in differential residential selections. Paper presented to the Symposium on Effects of Residential Mobility on the Wife. Indianapolis: University of Indiana-Purdue University Medical Centre.

Morrison, P. A. 1967. Duration of residence and prospective migration: the evaluation of a stochastic model. Demography 4:553-561.

Myers, George C., R. McGinnis, and G. Masrick. 1965. Preliminary assessment of a stochastic process model of internal migration. Population Index 31:256-7.

Rossi, H. 1955. Why Families Move: A Study in the Social Psychology of Urban Residential Mobility. Glencoe, Illinois: The Free Press.

Simmons, James W. 1974. Patterns of Residential Movement in Metropolitan Toronto. Toronto: University of Toronto Press.

_______. 1973. Net migration within metropolitan Toronto. In L. Bourne, R. MacKinnon and J. Simmons, (eds.). The Form of Cities in Central Canada: Selected Papers. Toronto: University of Toronto Press, Department of Geography Research Publication No. 11.

———, and J. Hardy. 1968. The Migration History Project: An interim report. Toronto: Ontario Institute for Studies in Education, Department of Education Planning.

Stone, Leroy O. 1971. Composition of Inter-City Migration Streams. Toronto: Ontario Institute for Studies in Education, Department of Educational Planning, Research Paper No. 1.

———. 1970. Migration in Canada. Regional Aspects. 1961 Census Monograph. Ottawa: Queen's Printer.

Received February, 1977; revised August, 1978.

Appendix: Expectations of Observed frequencies under Model Assumptions

Let $N_I(t)$ be the number of new families entering the population at time t in state i. We may calculate the expected number of families having any particular history over the period of study. Also define

$u_{ki}(s) = P_r\{\text{family is in state } i \text{ at } s + t \mid \text{the family entered the system at time } t \text{ in state } k\}$

$w_{ki}(s) = P_r\{\text{family enters state } i \text{ at } s + t \mid \text{the family entered the system at time } t \text{ in state } k\}$

Then for example, if we have only one year of observation, the expected number of families in i at time t_0 is

$$\sum_k \sum_{s=0}^{\infty} N_k(t_0 - s) u_{ki}(s).$$

We make an additional assumption,

$$N_k(t) = N_k \text{ for all } t$$

which we use throughout the remainder of the essay. Let $N = \sum_k N_k$.
Then

$$P\{\text{family starts in } k\} = N_k/N.$$

Note also

$$\begin{aligned}&\sum_{s=0}^{\infty} u_{ki}(s)\\ &= \sum_{s=0}^{\infty} P\{\text{family is in state } i \text{ at } s \mid \text{starts in } k \text{ at } 0\}\\ &(= E\{\sum_{s=0}^{\infty} \text{family is in state } i \text{ at } s \mid \text{starts in } k \text{ at } 0\})\\ &= E\{\text{amount of time a family spends in state } i \mid \text{starts in } k \text{ at } 0\}\end{aligned}$$

Then

$$\begin{aligned}\sum_{k,s} N_k u_{ki}(s) &= N \sum_k N_k/N \quad E\{\text{amount of time spent in } i \mid \text{family starts in } k\}\\ &= NG_i\end{aligned}$$

where G_i is the expected amount of time a family spends in state i, unconditional on starting state.

Theorem: For transient state i,

$$\sum_k \sum_{s=0}^{\infty} \frac{N_k}{N} w_{ki}(s) = \frac{1}{\mu_i} G_i.$$

We may call this quantity E_i, and interpret it as the expected number of transitions made into state i, unconditional on the starting state.

Proof:

$$u_{ki}(s) = \sum_{t=0}^{\infty} w_{ki}(t)\,\Phi_i(s\text{-}t)$$

$$\therefore \sum_{s=0}^{\infty} u_{ki}(s) = \sum_{s=0}^{\infty} \sum_{t=0}^{s} w_{ki}(t)\,\Phi_i(s\text{-}t)$$

$$= \sum_{t=0}^{\infty} \sum_{s=t}^{\infty} w_{ki}(t)\,\Phi_i(s\text{-}t)$$

$$= \sum_{t=0}^{\infty} w_{ki}(t) \sum_{s=0}^{\infty} \Phi_i(s'),\quad s' = s\text{-}t$$

$$= \sum_{t=0}^{\infty} w_{ki}(t)\,\mu_i$$

$$G_i = \sum_k \frac{N_k}{N} \sum_{s=0}^{\infty} u_{ki}(s)$$

$$= \mu_i \sum_k \sum_{t=0}^{\infty} \frac{N_k}{N} w_{ki}(t)$$

$$= \sum_k \sum_{t=0}^{\infty} \frac{N_k}{N} w_{ki}(t) = \frac{1}{\mu_i} G_i = E_i.$$

Since the theorem is only for transient i, there are no difficulties with either μ_i or E_i being infinite.

Armed with the above theorem, we can now compute expected numbers of observations of each type of history for the two year period of observations.

The expected value of a_i, the number of families observed initially (time t_o) in state i is

$$\sum_{t=-\infty}^{t_o} \sum_k N_k\,\mu_{ki}(t_o\text{-}t)$$

$$= \sum_{s=0}^{\infty} \sum_k N_k\,\mu_{ki}(s) = NG_i = NE\,\mu_i$$

The expected value of b_{ij}, the number of families observed initially (time t_o) in state i, and moving to j by time (t_o+1) is

$$\sum_{t=-\infty}^{t_o} \sum_k N_k \sum_{s=t}^{t_o} w_{ki}(s\text{-}t)\,P_{ij}\,\phi_{ij}(t_o+1\text{-}s)$$

$$= (\sum_k N_k \sum_{u=0}^{\infty} w_{ki}(\mu)) \sum_{r=1}^{\infty} p_{ij}\,\phi_{ij}(r)$$

$$= NG_i\,p_{ij}/\mu_i = NE_i\,p_{ij}$$