Working Paper 162

THE MEASUREMENT OF MIGRATION FROM CENSUS DATA AND OTHER SOURCES

P.H. Rees

Paper prepared for the Joint Meetings of the I.B.G. Population Geography and Quantitative Methods Study Groups on the Use and Analysis of Census Data, Sorby Hall, University of Sheffield, 23rd-25th September, 1976.

School of Geography
University of Leeds
Leeds LS2 9JT

September, 1976

CONTENTS

List of figures

ABSTRACT

In the first part of the paper a general classification of migration measures is introduced. The rates corresponding to these measures are defined together with the associated populations at risk. Then a particular problem in migration analysis is tackled. The questions asked in the British censuses of 1966 and 1971 about migration generate tabulations of migrants over the one year preceding the census date and over the five years preceding the census date. It is often observed that the numbers of migrants in the two periods are not linearly related; that is, the number of migrants over five years is less than five times the number over one year. A simple stochastic model embodying population accounting principles is developed to show why this is the case, and that the relationship between one year and five year figures is a complex one involving some multiple migrations, some return migrations, and some deaths.

1. The problem of meaning and measurement

Migration is a phenomenon that has captured the attention of social scientists for nearly a century and a vast body of literature has accumulated since Ravenstein's seminal papers (Ravenstein, 1885 and 1889). In most work on the subject, however, the meaning of the term "migration" is assumed to be known: little attention is paid to explicit definition of the method of measurement used and to the relation of this method of measurement to others. Even advanced texts reveal rather limited assumptions about the linkage of one measure of migration to another.*

In this paper an attempt will be made to remedy this lack of conceptual clarity in the discussion of migration, taking further some aspects of the work of Courgeau (1973), of Plessis-Fraissard (1975) and of Rees (1974). In particular, an attempt will be made to solve the puzzle that the number of migrants recorded over a five year interval is far less than five times the number recorded over a one year interval. Figure 1 shows, for example, the ratio of one year to five year migrants between the standard regions of Great Britain. These ratios vary from 0.251 for the North to East Midlands migrant flows to 0.497 for the Rest of the World to the North region flow. Or, in other words, the number of migrants recorded over five years in an interregional flow varies from 4 times to 2 times the number of migrants recorded over one year.

In the second section of the paper the various methods of measuring migration are outlined using migration history and time space diagrams. These are illustrated in section 3 using a population accounts table, and in section 4 the different kinds of migration rates associated with each type of count are described. Possible explanations for the five year-one year migration puzzle are then reviewed in section 5 of the paper. A possible process that links the two measures of migration is then described in section 6 of the paper in which data from the General Household Survey (hereafter referred to as GHS) on migration is used. Section 7 extends the analysis from a non-spatial to a multi-regional analysis

---

* Rogers (1975, p.70,71) reveals that "Crude estimates of annual data for the former (two-region migration data) were obtained by taking one-fifth of the published 5-year migration data " in computing sets of two region life tables for the U.S.A. The incorrectness of such an estimate will become apparent in this paper, though ironically, in an earlier work (Rogers, 1968), the author has already developed a technique to solve his problem (Chapter 4).

of the process. Some of the findings of this section can be employed to generate average 1 year migration from 5 year migration tables.

2. Alternative measures of migration

2.1 Initial definitions and classification

Migration can be defined as a permanent change of usual residence by a person or family or household. With Lee (1966) we place no restriction on the distance involved in such a relocation. It may be a move from one flat or apartment to another in the same building or it may be a move from a residence in one country to a home in another land. In the past, some researchers have drawn a distinction between moves between local communities (cities, labour markets) and moves within local communities. The former was labelled "migration", the latter "local mobility". This has been thought of as a distinction as well between moves of usual residence motivated by a relocation of employment and moves motivated by housing needs or wants. However, Harris and Clausen (1967) and Hyman and Gleave (1976) have shown that this distinction is false and that the placing of spatial constraints on the definition of migration should be avoided.

A migration is an event that happens to a person, family or household or an action taken, and it must be distinguished from the concept of migrant, a person who has in some specified period in the past experienced one or more migrations. This distinction between the event (migration) and the actor (migrant), made very clearly by Courgeau (1973), is a crucial one. A table of migrations involves counts of the number of migrations that have occurred irrespective, usually, of the number of migrants involved, and a table giving information about migrants involves counts of migrants irrespective of the number of migrations involved. Some data sources provide information on migrations (the International Passenger Survey as reported in O.P.C.S.,annual), some (the Migration Tables of the 1971 Census, O.P.C.S.,1974-5) give data on migrants only, while others (O.P.C.S., Social Survey Division, 1973) provide information on both. This kind of table is illustrated in Figure 2.

A. Five year migrants (1966-71)

Origin at start of period	N	YH	NW	EM	WM	EA	SE	SW	W	S
				Destination at end of period						
1 North (N)	-	32460	21880	14460	13250	6480	47530	10660	3850	14820
2 Yorks.& Humb. (YH)	36300	-	40220	40460	18760	10810	62570	17360	7030	11270
3 North West (NW)	18740	34850	-	19060	29580	8540	84270	27100	33620	14710
4 East Midlands (EM)	9570	35010	16280	-	25850	16180	57370	19320	6040	8450
5 West Midlands (WM)	10230	15230	32110	35660	-	8540	76350	40720	21120	9780
6 East Anglia (EA)	3390	7820	5220	12820	5510	-	53370	10020	3120	4030
7 South East (SE)	33720	47600	64890	73430	62750	109920	-	193920	36540	45140
8 South West (SW)	8010	10250	14930	12240	22840	9170	133950	-	14290	10800
9 Wales (W)	3610	5590	17790	6740	18450	3470	42790	20400	-	4340
10 Scotland (S)	16110	13600	21640	14150	13820	5870	64150	13260	3990	-
11 Rest of the World (RW)	29320	60330	85110	51900	78020	51060	525190	71070	21900	62380

B. One year migrants (1970-71)

1 North (N)	-	9290	7370	3630	3790	2190	15040	3870	1220	4910
2 Yorks.& Humb. (YH)	12790	-	13290	13190	5970	4320	21580	5610	2570	3470
3 North West (NW)	6790	10650	-	6320	9750	2980	29210	8950	10370	5200
4 East Midlands (EM)	4200	12170	5300	-	9610	6030	19110	6680	2330	3100
5 West Midlands (WM)	3550	5400	9750	12600	-	3040	25630	12590	6350	3340
6 East Anglia (EA)	1150	2510	1970	4490	2540	-	20460	3820	1170	1650
7 South East (SE)	12070	16570	22920	26390	21410	35410	-	66570	12050	7610
8 South West (SW)	2970	4600	4860	4230	8680	3390	49030	-	4730	4810
9 Wales (W)	1450	2040	6390	2420	5650	1280	13160	7000	-	1650
10 Scotland (S)	5540	4810	7220	4910	4630	2070	23790	5330	1700	-
11 Rest of the World (RW)	14570	19360	27330	16650	23140	20280	197220	29390	8840	28160

C. Ratio (Five year/One year)

1 North (N)	-	3.50	2.97	3.98	3.50	2.96	3.16	2.75	3.15	3.02
2 Yorks.& Humb. (YH)	2.84	-	3.03	3.07	3.14	2.50	2.90	3.10	2.73	3.25
3 North West (NW)	2.76	3.27	-	3.01	3.03	2.87	2.88	3.03	3.25	2.82
4 East Midlands (EM)	2.28	2.87	3.07	-	2.69	2.68	3.00	2.89	2.59	2.72
5 West Midlands (WM)	2.88	2.82	3.29	3.72	-	2.81	2.98	3.24	3.32	2.92
6 East Anglia (EA)	2.95	3.12	2.65	2.86	2.17	-	2.61	2.62	2.67	2.44
7 South East (SE)	2.79	2.87	2.83	2.79	2.79	2.93	-	2.92	3.03	2.62
8 South West (SW)	2.70	2.23	3.07	2.89	2.63	2.70	2.73	-	3.02	2.25
9 Wales (W)	2.49	2.74	2.79	2.79	3.27	2.71	3.25	2.92	-	2.63
10 Scotland (S)	2.91	2.82	2.99	2.88	2.99	2.83	2.70	2.49	2.35	-
11 Rest of the World (RW)	2.01	3.08	3.12	3.12	3.37	2.52	2.66	2.42	2.48	2.22

Source: O.P.C.S. (1974-75). Census 1971. England and Wales Migration Regional Reports. Part I. H.M.S.O. London and O.P.C.S. Edinburgh (1975) Census 1971. Scotland Migration Table Part I. H.M.S.O. Edinburgh.

Figure 1. Five-year migrants, one-year migrants and five year/one year ratios for interregional migration, 1971 Census.

Number of moves (migrations) made in past 5 years	Percent of all heads of household	Number of moves made of sample
0	64.6	0
1	23.4	2784
2	6.7	1594
3	3.2	1142
4	1.1	524
5 or over	1.0	714*
Total	100.0	6758
	N = 11899	

Source: O.P.C.S. Social Survey Division, 1973 GHS, Table 5.52 "Heads of household by number of moves in the past five years"

* An average of 6 moves assumed for this category.

Figure 2. Migrants and migrations, Great Britain, 1971

The tabulation (Figure 2) reveals that the 11,899 heads in the Great Britain sample made 6,758 migrations over the five years prior to 1971 or an average of 0.568 moves per person. However, only 35.4 percent of the heads (4212) can be classified as migrants and they made an average of 1.604 migrations.

With these definitions of migration and migrant in mind, we can classify the measures of migration provided in surveys or censuses. We distinguish five kinds of data source and associated migration or migrant indicator: migration histories; place of residence by place of birth tables; surviving migrant tables; other migrant types present in population accounts tables; and tables of migrations across significant boundaries. We also consider indirect measures of migration - not migration-a little later in this section of the paper. Examples of data items for individuals from each of these sources are illustrated in Figure 3.

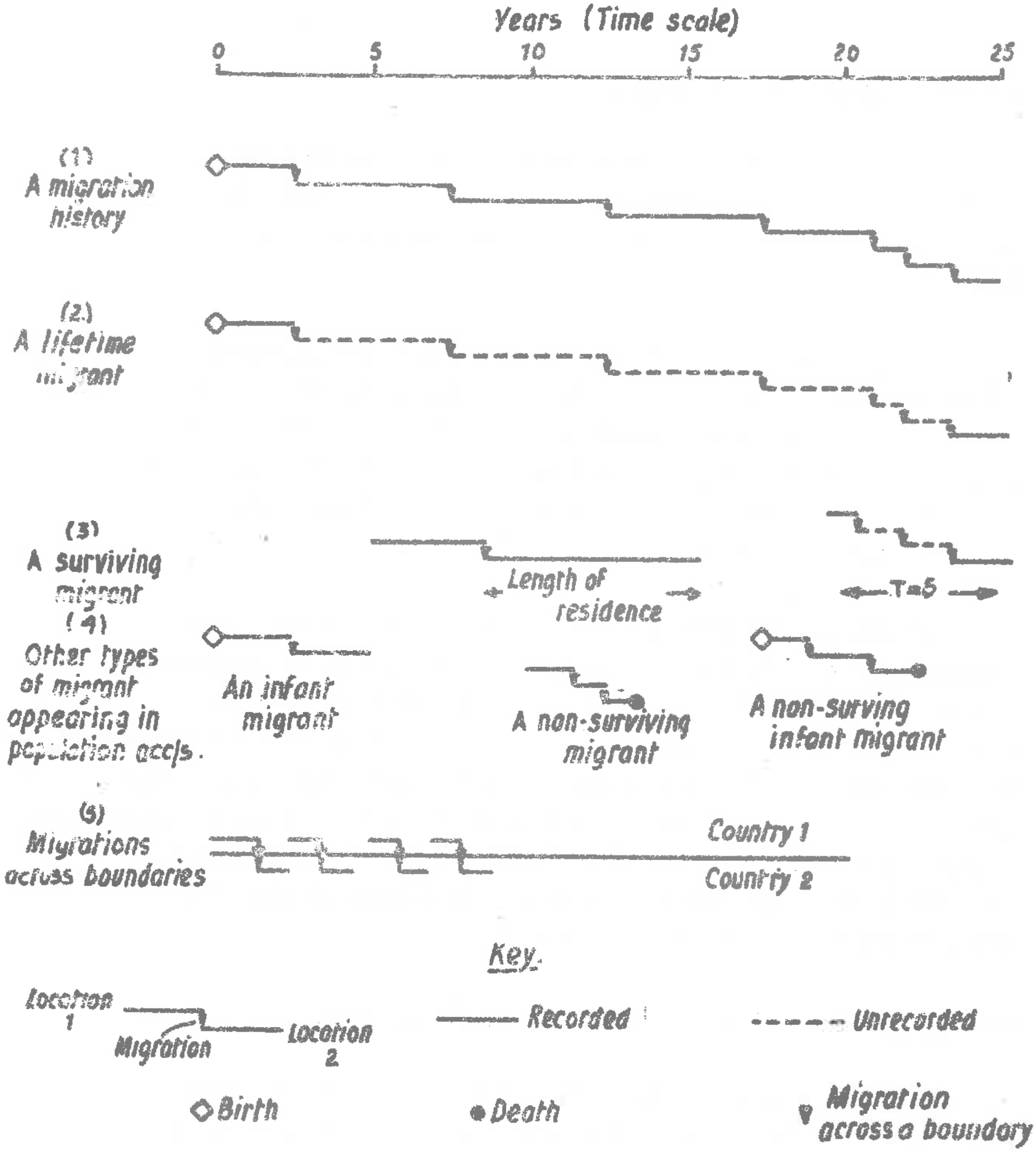


Figure 3 Examples of the different sources for migration indicators

2.2 Migration histories (type 1)

These are sets of life histories of individuals in which all residential locations occupied have been recorded. Most kinds of migration or migrant measure can be computed from information of this sort.

The migration history of the individual recorded in Figure 3 (1) is one which reveals that he made some seven migrations (the vertical drops with arrows attached) before the age of 25 and lived in some eight locations (horizontal extensions of the lifeline). He was a migrant in each quinquennium, but his migrant activity was greatest in years 20 to 25.

The sets of life histories available are usually partial in coverage either in time or in space or in coverage of the population except in countries such as Sweden in which population registers have a long history. Morrison and Relles (1974) has used the American metropolise Security Administration's Continuous Work History Sample to work out rates of in-migration and out-migration of the largest American metropolise. Plessis-Fraissard (1975) has used a sample of migration histories gathered in Leeds in 1973 to examine the dependence of migration rates on age and length of residence.

2.3 Place of residence by place of birth tables (type 2)

These are tables usually contained in a Census which cross classify persons recorded at their residences on census night by their places of birth (or the place of usual residence of their mother at time of maternity). The net relocation of persons between their birth date and the time of the census can be ascertained. The lifeline labelled (2) in Figure 3 illustrates the concept. The two known locations ("present" place of residence and place of birth) are shown as continuous lines and the intermediate and unknown locations are shown as pecked lines. Thus, lifetime migration tables underestimate the number of migrations experienced in an average lifetime.

Ravenstein (1885, 1889) in his seminal papers on "the laws of migration" made extensive use of such place of residence by place of birth tables produced for counties and boroughs in Great Britain and Ireland. More recently Friedlander and Roshier (1966) have attempted to use this data to infer the amount of decennial migration taking place (our type (3) data, that is). Rogers and Van Rabenau (1971) and Rogers (1975, pp. 172-185) have defined the mathematical model that makes possible the generation of intercensal migration estimates from sets of successive place of residence by place of birth tables.

2.4 Surviving migrant tables (type 3)

Many censuses now contain a question concerning the location of the persons enumerated in the census some T years ago. From the responses to this question can be determined the number of surviving migrants between places over the period t-T to t where t is the time of the census. Commonly, T takes on the value 1,2, 5 or 10 years. The information ascertained by a census at year 25 about the individual whose migration history is shown in Figure 3(1) is depicted in the rightmost lifeline in Figure 3(3). We know his initial location at t-5 or 20 years and his final location at t or 25 years but not the intermediate locations. In this case we know he was a migrant - that is, that he made at least one migration in the time interval.

Surviving migrant tables have been produced for the 1961 Census, the 1966 Sample Census and the 1971 Census in the U.K. and the information contained has been extensively analyzed by academics and used by planners to compute rates for forecasting purposes. Figure 1 shows the tables of migrants amongst standard regions in Great Britain over five years prior to census date in 1971 (Figure 1A) and over one year prior to census date (Figure 1B).

A second question sometimes asked in Censuses or in Surveys can also generate surviving migrant tables. This is the question which asks how long a person has been resident at his or her location on Census night. From this information the number of migrants within a specified period can be ascertained by cumulating the persons reporting lengths of last residence less than or equal to that time interval.

It might be thought that the omission of intermediate migrations in this kind of table was a serious deficiency. It is if we are interested in developing a time series of migrant tables, of course: five tables of surviving migrants over 1966-67, 1967-68, 1968-69, 1969-70 and 1970-71 would, of course, be preferable to the 1966-71 and 1970-71 tables actually available. But the absence of information on intermediate migrations should not be taken too seriously. Surviving migrant tables, in fact, supply exactly the right kind of transition data required in population accounts, for example, and in properly specified multi-regional population forecasting models (Rees and Wilson, 1976). It is the net relocation over an interval of time or the transition between states in a regional system that is required in accounting or forecasting models.

2.5 Other migrant types present in population accounts (type 4)

There are other kinds of migrants who are omitted in Census tables and in retrospective surveys but who are required for the construction of population accounts. Three types are important: surviving infant migrants, non-surviving migrants and non-surviving infant migrants.

Surviving infant migrants are children present at a point in time, born in the previous time interval, who have migrated (at least once) within it. The leftmost lifeline in Figure 3(4) shows such a migrant. Such migrants are normally omitted from surviving migrant tables - they are the "under 1's or under 5's" in the one year and five year tables respectively. Yet a census such as that of 1971 contains in its enumeration form the two questions necessary to generate infant migrant tables. These are the present residence question and the place of birth question. Infant migrants are lifetime migrants of an age less than or equal to the time interval of the retrospective migration question or the time interval used in population accounts. It would be relatively easy to generate such tables and it is hoped that this will be done for the 1981 U.K. Census.

Non-surviving migrants or migrants who die in a particular time interval of interest also figure in population accounts. These are persons alive at the beginning of the time interval and resident in a region of interest who migrate to another region and die there before the end of the time interval. It is possible to count up such migrants from full sets of migration histories (Figure 3(1)) but not from death records above as these give only the usual residence of the deceased immediately prior to death and not his or her usual residence one, two or five years ago. A simple model is used by Rees and Wilson (1976) to estimate the number of migrants who die from tabulations of surviving migrants and regional deaths.

The final migrant type under this fourth heading is the non-surviving infant migrant, a person born in the time interval of interest who migrates from one place to another within the time period and dies in the destination. The combination of birth, migration and death makes this kind of migrant rather rare. Again a population registration system would be needed to generate tables of non-surviving infant migrants. However, many local Child Health Departments in the U.K. make a practice of merging birth and infant death records in order to detect particular health problems, and these records, on a national basis, might yield the requisite statistics. Again a simple model is used to fill these missing items in population accounts tables.

2.6 Tables of migrations across significant boundaries (type 5)

All the data sources so far discussed have concerned persons and hence migrants rather than migrations with the exception of the first source. Information on international migration, except that which is contained in census tabulations, is recorded in the form of migrations (moves) across international boundaries by persons or families (see Jenkins, 1976 for a more extensive discussion). Country of last residence and nationality are usually ascertained by immigration authorities at the time of entry but there are very many problems involved in compiling comparable sets of statistics (see Jenkins, 1976). Figure 3(5) shows examples of lifelines crossing a border between two countries. Intention to reside for a specified period of time from a few months to a year in the country of entry is used to distinguish visitors from migrants, but it is not usually possible, except in countries with population registers to check whether the initially stated intention is carried through. It is not possible therefore to regard counts of migrations across borders as counts of the numbers of migrants of types (4) and (5). Migrants of each of these types can have made several migrations over the period.

It is also possible, in certain circumstances, for non-migrants to have made several migrations. Two situations in which this can occur are shown in Figure 4. In the first the person is classified as a non-migrant if we are interested only in looking at migration between regions i and j even though he or she has made two migrations internal to region i. In the second case the person has first migrated to region j and then returned to region i and is recorded there as a "surviving stayer". Individuals in either of these two situations would be recorded in the diagonals of the inter-regional migrant tables displayed in Figure 1,were numbers to be allocated to these cells.

2.7 Net migration or migrant tables (type 6)

The five data sources and associated migrant or migrations counts all involve direct measurements. That is, the people involved are asked questions about their migration behaviour. An alternative method is to use other statistics to infer the net balance in the number of migrations or migrants from population counts at successive censuses and from records of births and deaths in intervening years. This "residual" method of measuring migration derives from the "conservation of population" equation which can be stated as

$$P^i(t+T) = P^i(t) + B^i(t,t+T) - D^i(t,t+T) + M^{Ri}(t,t+T) - M^{iR}(t,t+T) \qquad (1)$$

where P refers to population, B to live births, D to deaths, M to migrations, i to the region of interest, R to the rest of the world,

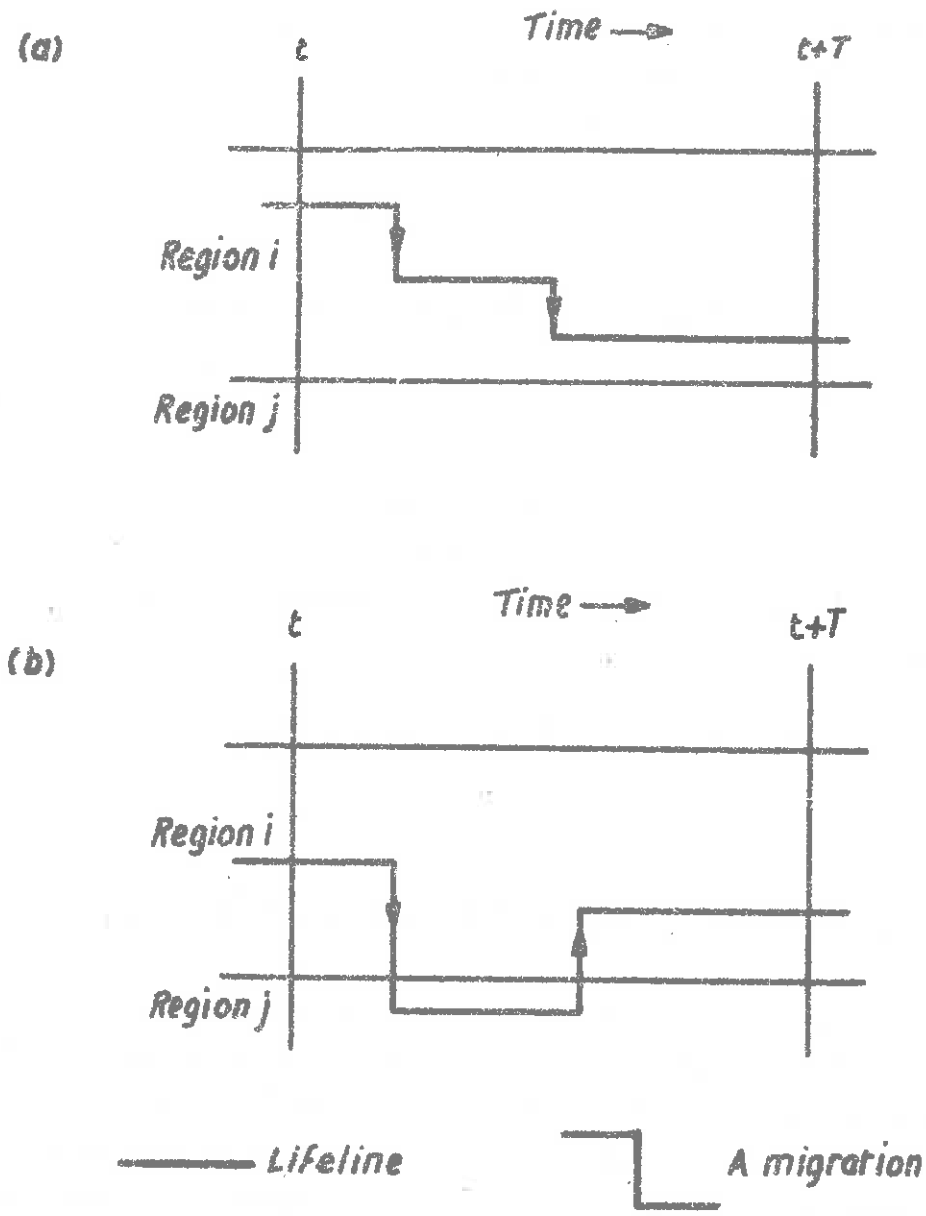


Figure 4 Two situations in which "non-migrants" make migrations

t to the start of the period, T to the length of the period, t+T to the end of the period and (t,t+T) labels the time period by its start and finish points.

We can rearrange equation (1) to read

$$M^{Ri}(t,t+T) - M^{iR}(t,t+T) = P^{i}(t+T) - P^{i}(t) - B^{i}(t,t+T) + D^{i}(t,t+T) \qquad (2)$$

where the items on the left hand side are unknown and on the right hand side are known. It is not generally possible to produce separate estimates of the two migration terms so that we usually write equation (2) as

$$N^{Ri}(t,t+T) = P^{i}(t+T) - P^{i}(t) - B^{i}(t,t+T) + D^{i}(t,t+T) \qquad (3)$$

where

$$N^{Ri}(t,t+T) = M^{Ri}(t,t+T) - M^{iR}(t,t+T) = - N^{iR}(t,t+T).$$

Note that, so far, we have written the equations in terms of migrations (the events) rather than migrants (the persons). We can rewrite equation (1) in terms of migrants of types (3) and (4) defined above. Let K refer to persons, i and R to the regions of interest and the rest of the world respectively as before; let ε refer to the "existence at the start of the period" life state, β to the "birth during the period" life state, δ to "death during the period life state" and σ to the "survival at the end of the period" life state. Using this notation a surviving migrant (type (3)) is denoted as $K^{\varepsilon(i)\sigma(R)}$, for example. A surviving infant migrant is represented as $K^{\beta(i)\sigma(R)}$; a non-surviving migrant is represented as $K^{\varepsilon(i)\delta(R)}$ and a non-surviving infant migrant as $K^{\beta(i)\delta(R)}$. Equation (1) now becomes

$$P^i(t+T) = P^i(t) + B^i(t+T) - D^i(t,t+T)$$

$$+ K^{\varepsilon(i)\sigma(R)}(t,t+T) - K^{\varepsilon(R)\sigma(i)}(t,t+T)$$

$$+ K^{\beta(i)\sigma(R)}(t,t+T) - K^{\beta(R)\sigma(i)}(t,t+T)$$

$$+ K^{\varepsilon(i)\delta(R)}(t,t+T) - K^{\varepsilon(R)\delta(i)}(t,t+T)$$

$$+ K^{\beta(i)\delta(R)}(t,t+T) - K^{\beta(R)\delta(i)}(t,t+T)$$

$$+ M^{iR}_{SUR}(t,t+T) - M^{Ri}_{SUR}(t,t+T) \qquad (4)$$

where M_{SUR} are the migrations surplus to those required to place the four migrant types (and non-migrants) in their respective categories. Thus, in Figure 4 the two migrations recorded in the lower diagram are surplus to place the person in the $K^{\varepsilon(i)\sigma(j)}$ category. The surplus concept was introduced in Rees (1974). If the two regions involved partition the world then the two surplus terms are equal, that is

$$M^{iR}_{SUR}(t,t+T) = M^{Ri}_{SUR}(t,t+T) \qquad (5)$$

and the last pair of terms can be omitted from the right hand side of equation (4).

We can similarly rearrange and shorten equation (4) to yield

$$\sum_{\alpha}\sum_{\omega} K^{\alpha(i)\omega(R)} - \sum_{\alpha}\sum_{\omega} K^{\alpha(R)\omega(i)} = P^i(t+T) - P^i(t)$$

$$- B^i(t,t+T) + D^i(t,t+T) \qquad (6)$$

where α represents a general label for the initial life state with two values ε and β, and ω represents a general label for the final life state with two values σ and δ. The most important point to note from this formal analysis of the residual method of estimating net migration is that it does not yield estimates of the number of surviving net migrations alone $(K^{\varepsilon(i)\sigma(R)}(t,t+T) - K^{\varepsilon(R)\sigma(i)}(t,t+T))$, as these are only one of four pairs of terms involved, but estimates of the net number all types ofmigrant.

3. An illustration of the various measures of migration

It is probably useful at this stage in the paper to illustrate with numbers the various measures of migration which have been outlined in section 2. For this purpose we use a population accounts table for the standard regions of Britain shown in Figure 5. The derivation of this table is described in Rees (1976) and the underlying theory is described in Rees and Wilson (1973, 1976). We will use the table in a later section to illustrate the various ways in which migration rates can be measured.

The Figure 5 accounts are for the one year period prior to the Census of 1971 (April 25/26). The rows represent the origin states - either existence at "census date" 1970 or birth in the one year prior to the census in one of the standard regions of England, in Wales, in Scotland, or in the rest of the World. The columns represent the destination states - either survival in one of the regions at the time of the census or death there in the year preceding. Only a small portion of the migration histories of persons (some 53,978,535 people) present in Great Britain at census date 1971 is utilized in the table (only one year's worth). Lifetime migrants under 1 year of age are represented by the off-diagonal elements of the bottom left quadrant of the table where the distribution of infant residents of the regions by place of birth has been estimated. Surviving migrants are contained in the off-diagonal portion of the existence-survival quadrant. We have already seen these numbers in Figure 1B. Non-surviving migrants are contained in the off-diagonal portion of the existence-death quadrant of the table, again estimated rather than measured. Finally, non-surviving infant migrants are shown in the bottom right hand quadrant of the table although these are estimated to be only 41 in number. By combining the numbers from each of the quadrants we can obtain a minimum estimate of the number of migrations that occurred amongst the regions to which should be added, where available, the further migrations surplus to placing persons in their final positions in the table.

The number of net migrants of each type between pairs of regions can only be computed by assembling the appropriate figures from Figure 5. Thus, the number of net migrants between East Anglia and the

Initial State \ Final State	1 N	2 YH	3 NW	4 EM	5 WM	6 EA	7 SE	8 SW	9 W	10 S	11 RW
	Survival at C.D. 1971										
Existence at C.D. 1970											
1. North	3180252	9290	7370	3630	3790	2190	15040	3970	1220	4910	25275
2. Yorkshire & Humberside	12790	4633103	13290	13190	5970	4320	21580	5610	2570	3470	27806
3. North West	6790	10650	6530593	6320	9750	2980	29210	8950	10370	5200	43952
4. East Midlands	4200	12170	7300	3237452	9610	6030	19110	6660	2330	3100	25545
5. West Midlands	3550	5400	9750	12600	4930317	3040	25630	12590	6350	3340	27877
6. East Anglia	1150	2510	1970	4490	2540	1554014	20460	3820	1170	1650	28194
7. South East	12070	16570	22920	26330	21410	35410	16551698	66570	12050	17610	193299
8. South West	2970	4600	4260	4230	8680	3390	49050	3553826	4730	4810	35450
9. Wales	1450	2040	6390	2420	5650	1280	13160	7000	2631000	1650	10185
10. Scotland	5540	4810	7220	4910	4630	2070	23790	5330	1700	5223725	57612
11. Rest of the world	14570	19630	27330	16650	23140	20280	197220	29390	8840	28160	0
Sub totals	5245362	4725773	6545393	3532312	5025037	1635014	16965925	3703656	2632330	5136154	475275
Birth in 1970-71											
1. North	52151	73	58	29	50	17	119	31	10	39	200
2. Yorkshire & Humberside	109	80569	113	112	51	37	184	48	22	30	238
3. North West	57	90	112265	53	82	25	246	75	87	44	370
4. East Midlands	35	103	45	55667	81	51	161	56	20	26	216
5. West Midlands	31	47	85	110	88686	26	223	110	55	29	243
6. East Anglia	9	20	16	36	20	25873	163	30	9	13	224
7. South East	94	129	179	206	167	277	265676	520	94	139	1510
8. South West	23	35	37	32	66	26	376	56279	36	37	272
9. Wales	11	16	51	19	45	10	104	55	42533	13	81
10. Scotland	47	41	61	42	39	18	201	45	14	86708	488
11. Rest of the world	256	344	479	292	406	356	3459	515	155	494	0
Sub totals	52823	81467	113387	56593	87673	26716	270912	57764	43035	87571	3842
TOTALS	5293185	4807240	6759380	3588910	5115760	1661730	17236840	3761400	2725365	5223725	479117

Initial State \ Final State	Sub totals	1 N	2 YH	3 NW	4 EM	5 WM	6 EA	7 SE	8 SW	9 W	10 S	11 RW	D.D. Totals	TOTALS
		Death in 1970-71												
Existence at C.D. 1970														
1. North	3256867	38271	55	45	20	19	12	83	24	8	29	147	38714	3295581
2. Yorkshire & Humberside	4748773	74	55031	63	72	31	24	119	34	16	21	165	56470	[illegible]
3. North West	6673765	39	63	82920	35	50	17	[illegible]	55	65	31	271	84710	6757475
4. East Midlands	3331557	24	72	33	36111	49	33	106	41	15	18	140	[illegible]	3368193
5. West Midlands	5041044	21	32	61	69	52030	17	142	77	40	20	143	52622	5093666
6. East Anglia	1621973	7	15	12	25	13	17392	113	23	7	16	156	18265	1640341
7. South East	16975997	70	98	143	145	110	196	187508	409	76	105	1069	189929	17165925
8. South West	3676576	17	27	30	23	44	19	271	45000	30	27	218	45708	3722284
9. Wales	2682225	8	12	40	13	29	7	73	43	33507	10	64	33906	2716131
10. Scotland	5179566	32	28	45	27	24	11	132	33	11	60633	342	61318	5241184
11. Rest of the world	385210	65	116	170	91	118	112	1090	181	55	167	0	2165	387395
Sub totals	53513664	39648	53349	83563	36631	52487	18430	189798	45020	33930	60526	2718	613567	[illegible]
Birth in 1970-71														
1. North	52757	305	0	0	0	0	0	0	0	0	0	1	306	[illegible]
2. Yorkshire & Humberside	81513	0	480	0	0	0	0	1	0	0	0	1	482	[illegible]
3. North West	113392	0	0	703	0	0	0	1	0	0	0	1	705	114097
4. East Midlands	56461	0	0	0	308	0	0	0	0	0	0	1	309	56770
5. West Midlands	89645	0	0	0	0	457	0	1	0	0	0	1	459	90104
6. East Anglia	26413	0	0	0	0	0	145	0	0	0	0	1	146	26559
7. South East	268990	0	0	1	1	0	1	1478	2	0	0	4	1487	270477
8. South West	57219	0	0	0	0	0	0	1	349	0	0	1	351	57570
9. Wales	42938	0	0	0	0	0	0	0	0	268	0	0	268	43206
10. Scotland	87704	0	0	0	0	0	0	1	0	0	518	1	520	88224
11. Rest of the world	6756	1	1	1	1	1	1	10	2	0	1	0	19	6775
Sub totals	883788	306	481	705	310	458	147	1493	353	268	519	12	5052	[illegible]
TOTALS	54457652	39954	56830	84268	36941	52945	18577	191291	46273	34198	61592	2730	624519	[illegible]

Figure 5. Local [illegible] accounts for British regions, 1970-71 (adjusted to [illegible] 1971 populations).

South East is as follows:

$$
\begin{aligned}
N^{67} &= (K^{\varepsilon(6)\sigma(7)} - K^{\varepsilon(7)\sigma(6)}) \\
&+ (K^{\beta(6)\sigma(7)} - K^{\beta(7)\sigma(6)}) \\
&+ (K^{\varepsilon(6)\delta(7)} - K^{\varepsilon(7)\delta(6)}) \\
&+ (K^{\beta(6)\delta(7)} - K^{\beta(7)\delta(6)}) \\
&= (20460 - 35410) \\
&+ (\ \ 163 - \ \ 277) \\
&+ (\ \ 113 - \ \ 196) \\
&+ (\ \ \ \ 0 - \ \ \ \ 1) \\
&= (-14950) + (-114) + (-83) + (-1) \\
&= -15148
\end{aligned} \tag{7}
$$

There is a net flow of 15148 migrants from the South East into East Anglia.

Total net migrants for East Anglia, say, can be computed in a similar fashion by working out the appropriate in-and out-migrant totals:

$$
\begin{aligned}
N^{*6} &= (\sum_{j\neq 6} K^{\varepsilon(j)\sigma(6)} - \sum_{j\neq 6} K^{\varepsilon(6)\sigma(j)}) \\
&+ (\sum_{j\neq 6} K^{\beta(j)\sigma(6)} - \sum_{j\neq 6} K^{\beta(6)\sigma(j)}) \\
&+ (\sum_{j\neq 6} K^{\varepsilon(j)\delta(6)} - \sum_{j\neq 6} K^{\varepsilon(6)\delta(j)}) \\
&+ (\sum_{j\neq 6} K^{\beta(j)\delta(6)} - \sum_{j\neq 6} K^{\beta(6)\delta(j)})
\end{aligned} \tag{8}
$$

where the asterisk, *, represents summation over all regions. Equation (8) can be evaluated numerically as

$$N^{*6} = (80990 - 67954)$$
$$+ (\ \ 843 - \ \ 540)$$
$$+ (\ \ 448 - \ \ 381)$$
$$+ (\ \ \ \ 2 - \ \ \ \ 1)$$
$$= (13036) + (303) + (67) + (1) = 13407 \quad (9)$$

An alternative way of computing this net migrant total is to use the residual method discussed in section 2.7. We employ the population, birth and death totals given on the margins of the table:

$$N^{*6} = P^6(\text{c.d. } 1971) - P^6\ (\text{c.d. } 1970) - B^6(1970,1971) + D^6(1970,1971)$$
$$= 1661730 - 1640341 - 26559 + 18577$$
$$= 13407 \quad (10)$$

This illustrates numerically the point that the net migrant balance calculated by the residual method refers to all four types of migrant involved in population accounts.

4. Rates of migration and associated populations at risk

The measures of migration described in section 2 and illustrated in section 3 are counts of persons or events. It is clearly useful in many instances to convert these counts into rates in order to measure the relative speed at which the migration process is operating. Migration rates are used when the characteristics of migrants are studied and some of the rates are employed in forecasting models.

The problems of proper rate and probability definition and proper specification of the populations at risk in the case of full migration histories in relation to lengths of residence and age have been outlined very effectively by Courgeau (1973) and Plessis-Fraissard (1975). In those papers the importance of being clear about the time period (prospective or retrospective) involved in the measurement of the migration events was stressed. Rates are rarely measured from place of residence by place of birth tables as the population at risk of experiencing the event (lifetime migration) are a mixture of the region-of-birth cohorts stretching back 100 years into the past.

The migrant types involved in population accounts tables can be considered together. If a prospective or forward looking point of view is adopted then transition rates are those most appropriately defined. A population flow is divided by its row total to yield a rate of transition from the row state to the column state. If we define h to be such a transition rate then typical rates are computed thus:

$$h^{\varepsilon(i)\sigma(j)} = K^{\varepsilon(i)\sigma(j)} / K^{\varepsilon(i)*(*)}$$

$$h^{\varepsilon(i)\delta(j)} = K^{\varepsilon(i)\delta(j)} / K^{\varepsilon(i)*(*)}$$

$$h^{\beta(i)\sigma(j)} = K^{\beta(i)\sigma(j)} / K^{\beta(i)*(*)}$$

$$h^{\beta(i)\delta(j)} = K^{\beta(i)\delta(j)} / K^{\beta(i)*(*)} \qquad (11)$$

where $K^{\varepsilon(i)*(*)}$ is the start of period population of region i and $K^{\beta(i)*(*)}$ is the total of live births in region i in the period. A full matrix of h rates for the British regional accounts is given in Rees (1976). Figure 6 shows the transition rates calculated for the East Midlands region. The rates in the first row of the table give the chances of persons located initially in the East Midlands of surviving and being relocated in other regions after a year (columns 1,2,3,5 to 11). Column 4 shows the rate at which persons can expect to survive and stay in the East Midlands. The total at the end of the row shows the rate at which people can expect to survive anywhere after one year has elapsed. The second row of the table shows the rates of migration and non-survival and in the case of column 4 of staying and non-survival: the total at the end of the row is the rate of dying anywhere. If we combine the totals for the rates in the first two rows we obtain a grand total of 1 indicating that these transition rates are genuine probabilities. Similar rates are calculated in the third and fourth rows: these are

	North	Y.& Humb	N.West	E.Mid.	W.Mid.	E.Ang.	S. East	S.West	Wales	Scotland	Rest of the World	Total
	1	2	3	4	5	6	7	8	9	10	11	
ES rates	.001247	.003613	.001574	.961191	.002853	.001790	.005674	.001983	.000692	.000920	.007584	.989121
ED rates	.000007	.000021	.000010	.010721	.000015	.000010	.000031	.000012	.000004	.000005	.000042	.010879
												1.000000
BS rates	.000617	.001814	.000793	.980571	.001427	.000898	.002836	.000986	.000352	.000438	.003805	.994557
BD rates	.0	.0	.0	.005425	.0	.0	.0	.0	.0	.0	.000018	.005443
												1.000000

ES - Existence and survival ED - Existence and death BS - Birth and survival BD - Birth and death

Figure 6 Transition rates for the East Midlands region, 1970-71

	North	Y.& Humb.	N.West	E.Mid.	W.Mid.	E.Ang.	S.East	S.West	Wales	Scotland	Rest of the World	Total
	1	2	3	4	5	6	7	8	9	10	11	
ES rates	.001071	.003892	.001865	.955317	.003718	.001325	.007787	.001248	.000714	.001449	.004913	.983299
BS rates	.000009	.000033	.000016	.016426	.000032	.000011	.000061	.000009	.000006	.000012	.000086	.016701
												1.000000
ED rates	.000541	.001949	.000947	.977532	.001868	.000677	.003925	.000623	.000352	.000731	.002463	.991608
BD rates	.0	.0	.0	.008338	.0	.0	.000027	.0	.0	.0	.000027	.008392
												1.000000

ES - Existence and survival ED - Existence and death BS - Birth and survival BD - Birth and death

Figure 7 Admission rates for the East Midlands region, 1970-71 the probabilities of ~~survival~~ and of relocation and staying and of relocation and staying and of death, given birth in the period.

Each of these rates of transition can be expressed as the product of a probability and a conditional probability. For example,

$$h^{\varepsilon(i)\sigma(j)} = h^{\varepsilon(i)\sigma(*)} \cdot P(j/\varepsilon(i),\sigma(*)) \qquad (12)$$

that is the rate of migrating from region i to region j and surviving there is equal to the probability of surviving anywhere given you started in region i, and the conditional probability of being placed in region j at the end of the period given that you started in region i and survived. The probability of migrating from the East Midlands to the South East is, using equation (12)

$$h^{\varepsilon(4)\sigma(7)} = (.989121)\ (.005736) = (.005674) \qquad (13)$$

Similar conditional probabilities can be defined for the other transition types.

Transition rates are used in "push" type models of population change. Sometimes, however, size of the final population or of the immigrant total may be fixed. In this situation size of admission rates are more appriориate. These are the rates at which the population admits persons from the various sources and they are formally defined as

$$a^{\varepsilon(i)\sigma(j)} = K^{\varepsilon(i)\sigma(j)} / K^{*(*)\sigma(j)} \qquad (14)$$

$$a^{\beta(i)\sigma(j)} = K^{\beta(i)\sigma(j)} / K^{*(*)\sigma(j)} \qquad (15)$$

$$a^{\varepsilon(i)\delta(j)} = K^{\varepsilon(i)\delta(j)} / K^{*(*)\delta(j)} \qquad (16)$$

$$a^{\beta(i)\delta(j)} = K^{\beta(i)\delta(j)} / K^{*(*)\delta(j)} \qquad (17)$$

where $K^{*(*)\sigma(j)}$ is the final population of the region at the end of the period and $K^{*(*)\delta(j)}$ the total of deaths during the time interval.

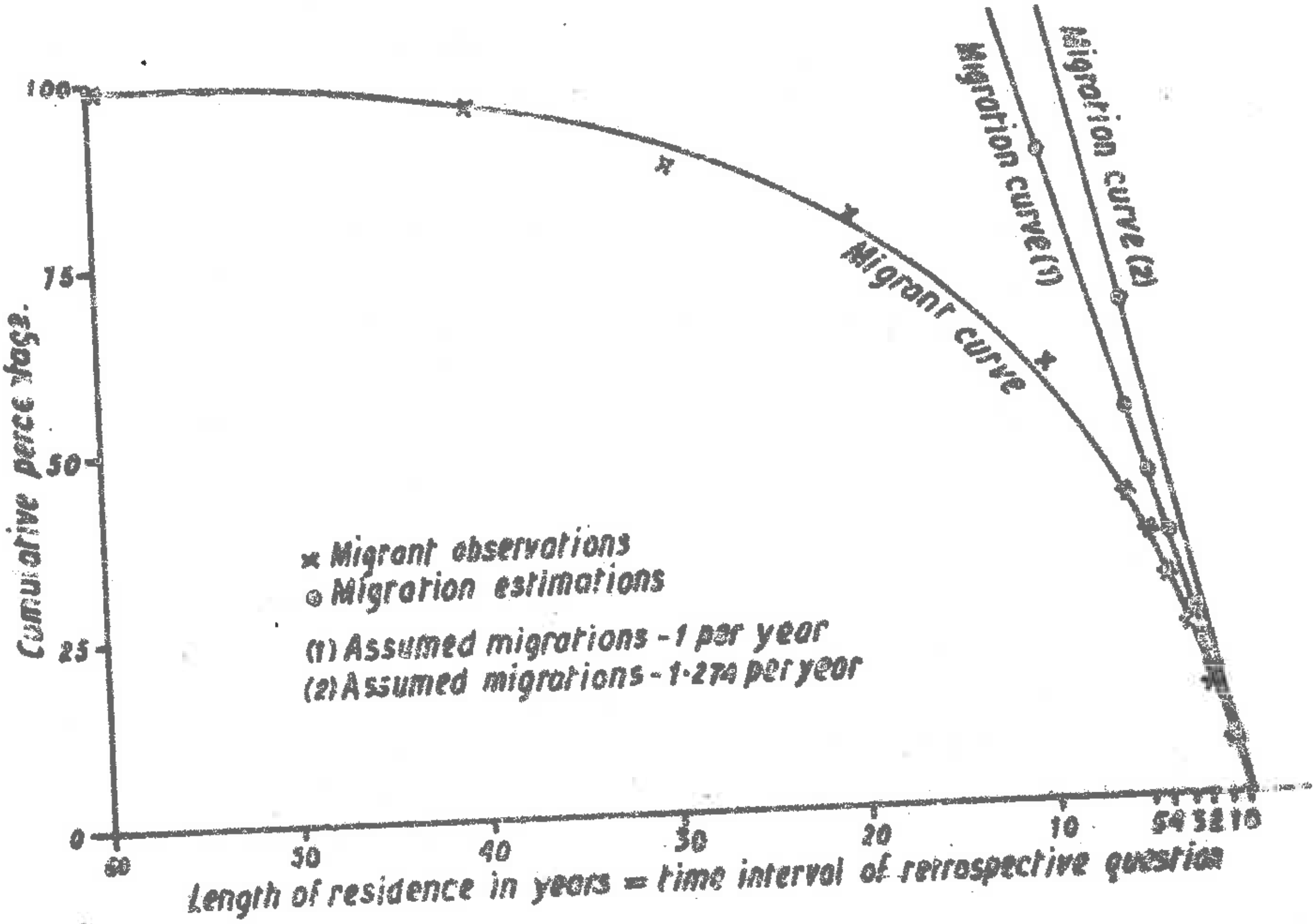
These rates again sum to 1 and can be regarded as probabilities Similar total and conditional rates can be defined. Admission rates for the East Midlands region are shown in Figure 7.

Most kinds of migrant data can be used to generate probabilistic rates. There is no guarantee, however, that rates involved counts of migrations will add up to 1 in the same way. This is because we are dealing with events rather than person transitions. The proper population at risk for division into a count of events is a sum of all the elements in the accounts table weighted by the risk of persons in that element of experiencing that event. Most weights are so small that they can conveniently be set to zero. Persons at risk of migrating from region i to region j in a time interval include all those who start the period in region i. Those who stay and survive there will be exposed all the period and are assigned a weight of 1; those who migrate to another region k and survive from i are assigned a weight of a ½, and so on. More details of the process of estimating multi-regional populations at risk are given in Rees and Wilson (1976) and Rees (1975), where they are applied in the context of deaths and births. They apply equally, though in a more complicated fashion, to migrations, and might be used, for example, in calculating international migration rates (Jenkins, 1976).

Having reviewed in fairly general terms the concepts involved in counting and rating migrants and migrations we now turn to the particular conceptual problem posed earlier in the paper and exemplified in Figure 1.

5. Possible explanations for the differences between the number of five year migrants and five times the number of one year migrants

In Figure 1 we examined the ratio of five year to one year migrants. A more extensive view of the relationship between the number of migrants counted and the length of the period over which they are counted is provided by tables constructed from a question about the length of residence at the preset address. Figure 8 presents a tabular distribution of responses by heads of households in the 1971 General Household Survey (O.P.C.S. Social Survey Division, 1973). If these are cumulated (as in the second column of the table contained in Figure 8) the cumulative percentages



…th of …idence (1)	% (2)	Time interval assumed (3)	Cumulative % (4)	Multiple of under 1 figure (5)	Estimated number of migrations per person At 1 per year (6)	At 1.274 per year (7)
…er 1	8.7	1	8.7	1.00	.087	.111
…ut under 2	6.2	2	14.9	1.71	.174	.222
…ut under 3	8.0	3	22.9	2.63	.261	.332
…ut under 4	6.8	4	29.7	3.41	.348	.443
…ut under 5	6.2	5	35.9	4.13	.435	.554
…ut under 6	4.6	6	40.5	4.66	.522	.665
…ut under 10	18.3	10	58.8	6.76	.870	1.108
…ut under 20	20.5	20	79.3	9.11	1.740	2.216
…ut under 30	8.0	30	87.3	10.03	2.610	3.324
…ut under 40	8.5	40	95.8	11.01	3.480	4.432
…and over	4.2	60	100.0	11.49	5.220	6.648
	100.0					

…re for column (2) data: O.P.C.S. Social Survey Division (1973) GHS. Table 5.30 Heads of household by age by length of residence at present address. Survey N = 11944.

…gure 8 The relationship between number of migrants and length of the time interval of the migration question represent the proportion of heads who have moved at least once in the years prior to the survey indicated by the rightmost digit of the length of residence category. Thus, 8.7 percent of heads are recorded as one year migrants and some 35.9 percent as five year migrants. The ratio of one to the other is 4.13, higher than in the Figure 1c off-diagonal elements. The survey includes all migrations undertaken by heads of households; Figure 1's off-diagonal elements involve only inter-regional migration.

We can also attempt an estimate of the number of migrations made by the migrants defined in column (4) of the bottom half of Figure 8. In column (6) we have assumed that only 1 migration per year per person is made and that this number is constant backwards over time. Thus, we have a linear increase of the migration rate over time. However, we can work out from Table 5.52 of the GHS (O.P.C.S. Social Survey Division, 1973) that heads of households make .5 4 of a migration in five years, an annual average of 1.274 per migrant and column (7) shows this rate linearly extrapolated

If we plot the cumulative percentages against the length of residence (interpreted as the time interval in a retrospective question on migration) we obtain the function shown at the top of Figure 8. This function is one which shows a tendency to decline in rate of ascent with an increase of age: fewer and fewer migrants are added as the time interval increase. A fairly simple process can be suggested to account for the form of the function.

Let us assume that the propensity of persons to become migrants over one year has remained constant over the time spans represented by the migration histories reported in the table. Or rather let us assume that the admission rate has remained constant at 8.7 percent or .087. Going backwards in time one year one has 1-.087 or .913 of the original sample of heads who are non-migrants. If we assume that .087 of these become migrants in the next year back, this leaves .913 - (.913) (.087) = .834 as non-migrants after two years and therefore 1-.834 = .166 as migrants after two years. We can repeat the process for as many years back as we wish.

We can generalize this process as follows. Let $m(1)$ be migrant admission rate for a one year period and $n(1)$ be the non-migrant admission rate or the proportion of the population or sample who are still non-migrants after one year. Then we can say that

$$n(T) = n(1)^T$$

$$m(T) = 1 - n(1)^T$$

$$\text{or} \quad m(T) = 1 - (1 - m(1))^T \tag{18}$$

Figure 9 compares the predictions of this simple model with those observed in Figure 8. The fit of model and observations is quite close particularly for the 5 year time interval. For time intervals up to 10 years the model tends to overpredict the proportion of heads who are classified as migrants and then to underpredict for periods of 20 and 30 years in length. For the last category in Figure 8, 40 years and over length of residence we have assumed a limit of 60 years since beyond that length of time very few people will be in a position to respond.

The same analysis can be applied on a regional basis to the inter-regional migrant tables given earlier in Figure 1. We can compute a rate at which migrants are admitted to the census date 1971 population for one year prior to the census and for five years prior to the census using migrant figures from Figure 1 and population figures from Figure 5. The observed five year rates can then be compared with those generated by equation (18) using the observed one year rates. Again the assumption is made that the one year rates for 1970-71 also characterize the other four years of the 1966-71 intercensal period. The rates are displayed in Figure 10, in the leftmost three columns. The model clearly overestimates the proportion of persons living in each of the regions who will be migrants by from 18 to 85 percent.* Clearly, equation (18) is an inadequate representation of the process at work.

---

* Percent = [(Observed-Model)/Observed] x 100

Length of residence (up to) or time interval in retrospective question T	Observed m(T) from Figure 8	Model m(T) derived from $1-(1-m(1))^T$ (m(1)=.087)	Difference (Observed -Model)
2	.149	.166	-.017
3	.229	.239	-.010
4	.297	.305	-.008
5	.359	.366	-.007
6	.405	.421	-.016
10	.588	.598	-.010
20	.873	.838	+.035
30	.958	.935	+.023
60*	1.000	.996	+.004

* An estimated T only.

Figure 9. Observed and model migration rates, GHS sample

	Total in-mig. rate: one year im(1)	Total in-mig. rate: five yr. im(5)	Model im(5) $1-(1-m(1))^5$	Total mig. rate: one yr. m(1)	Total mig. rateL Five yr. m(5)	Model m(5)
North	.0197	.0512	.0948	.1101	.3385	.4420
Yorks & Humb.	.0182	.0547	.0877	.1061	.3382	.4197
North West	.0157	.0474	.0763	.1056	.3233	.4276
East Midlands	.0280	.0829	.1323	.1080	.3254	.4354
West Midlands	.0186	.0565	.0892	.1074	.3302	.4334
East Anglia	.0487	.1384	.2211	.1280	.3519	.4958
South East	.0240	.0666	.1144	.1269	.3584	.4927
South West	.0398	.1555	.1839	.1266	.3602	.4916
Wales	.0188	.0556	.0907	.0938	.2913	.3889
Scotland	.0122	.0357	.0597	.1187	.3504	.4683

im - in-migrant rate

m - migrant rate

Figure 10 Observed and model migration rates, British regions

The rates on the left hand side of Figure 10 refer to migrants moving into a region from elsewhere. They do not include migrants internal to the region. These are persons present in the region at census date 1971 and at census date one year or five years previously but who moved home within the region. If we add these to the in-migrant figures we obtain a total of all persons who have made at least one migration in the one year or five year period. This total is more comparable to the sample population from which the observed m(T) rates were computed in Figure 9. If we apply the equation (18) model in this context we obtain the results shown on the right hand side of Figure 10. The absolute differences between observed m(5) rate and model m(5) rate are larger than with in-migrants above but the relative discrepancy is much smaller ranging from 24-41 percent overprediction.

It is difficult to put forward convincing explanations for the differences between Figure 9's results and those of Figure 10. The populations to which the two sets of results refer do differ but hardly enough to explain the differences in model performance. Figure 9 refers to heads of household, Figure 10 to all persons aged 1 and over or 5 and over. Both cover Great Britain. Heads of household are less migratory than the whole population over one year but slightly more migratory over five years (m(1) = .087 for GHS heads, .116 for the census population; m(5) = .359 for GHS heads and .342 for the census population). It may be that in the general population persons who have not migrated over a period of time are less likely to migrate than those who have already migrated. In this context a sort of "cumulative inertia" in reverse may be present although Plessis-Fraissard (1975) has thrown doubt on the concept where rates and populations at risk are correctly defined. It is clear from a comparison of Figure 9 and the lefthand side of Figure 10 that a more complex process than that of equation (18) is involved when an interregional framework is employed.

Let us review the possible explanations for the differences between the migrant and migrations curves shown in Figure 8 in the light of what we have learnt about the measurement of migration. Here we will take a prospective point of view and look forward one year and five years rather than backward. We can distinguish three contributions to the numerical difference between migrations and migrants: multiple migration, return migration (a particular kind of multiple migration) and the migrations of migrants who die. When migrants make many migrations the count of migrations increases while that of migrants remains static. The difference in counts we shall call "surplus" migrations. The number of surviving migrants may be decreased by death when the migrant moves into the non-surviving category or by return migration when the migrant moves back into the stayer category by returning to his or her original region. Figure 11 illustrates these effects. After one year the three persons are all classified as surviving migrants and have made 3 migrations. By the time five years have elapsed we have only 1 surviving migrant, 1 stayer and a total of 10 migrations. The surplus migrations count is the difference between the total of migrants (surviving and non-surviving) and the total of migrations. Of course, Figure 11 rather exagerates the relative importance of return migration, migration and death and their combination return migration and death. We will make estimates of the relative contributions of these components in the sections that follow. We now describe a simple probability model, extending that of equation (18), but in a prospective direction.

6. A simple model of the migration process

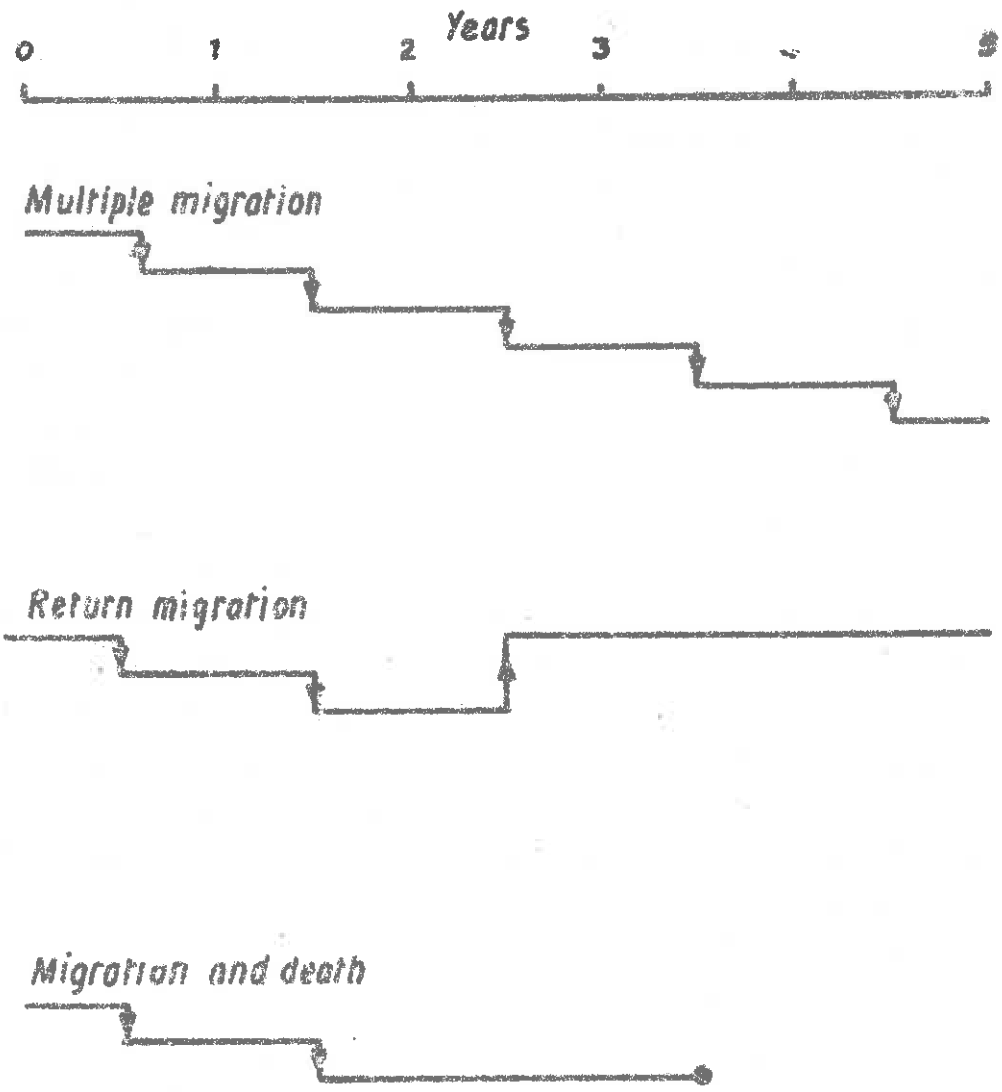
We will model migration here as a simple branching process (see Figure 13) in which four branches or outcomes are recognized in each year. These are staying (in the same location) and surviving, migrating and surviving, staying and dying, and migrating and dying. These are the four event sequences faced by an individual at the beginning of the year. We need to estimate the probability of each event sequence. To do this we form the probability matrix shown in Figure 12.

To fill in the matrix we use four known probabilities. We know the probability of having made at least one migration over the year prior to the 1971 GHS (O.P.C.S. Social Survey Division, (1973, Table 5.50) as .087. This is conditional on having survived the year so that we can set

$$P(M/S) = .087 \tag{19}$$

and

$$P(R/S) = .913 \tag{20}$$



Number up to year :	1	2	3	4	5
Surviving stayer	0	0	1	1	1
Surviving migrants	3	3	2	1	1
Non-surviving migrants	0	0	0	1	1
Migrations	3	6	8	9	10
"Surplus" migrations	0	3	6	7	8

Figure 11 Multiple migration, return migration and migration and death

Notation: P = probability

		Surviving S	Dying D	Totals
Staying or Remaining	R	P(R,S)	P(R,D)	P(R)
Migrating	M	P(M,S)	P(M,D)	P(M)
Totals		P(S)	P(D)	P(*)

Numerical values for heads of households in 1971 GHS

		Surviving S	Dying D	Totals
Staying	R	.891	.023	.914
Migrating	M	.085	.001	.086
Totals		.976	.024	1.000

Figure 12. The probability matrix for event sequences over one year

It is also possible to work out an approximate death rate for heads of household either by applying death probabilities for 1970-71 aggregated to persons from the separate rates for each sex taken from Appendix Table V of O.P.C.S. (1972) to the sample distribution of heads by age or by applying the sex-disaggregated probabilities to an estimated distribution at census date 1971 of heads of households by age and sex. The latter method is used here and yields an estimated death probability of .024. Thus, we assume that

$$P(D) = .024 \tag{21}$$

and

$$P(S) = .976 \tag{22}$$

Using the usual probability laws we can now estimate P(R,S) and P(M,S)

$$P(R,S) = P(S).\ P(R/S) = .976 \times .913 = .891 \tag{23}$$

$$P(M,S) = P(S).\ P(M/S) = .976 \times .087 = .085 \qquad (23)$$

For persons who die we assume that their migration probability will be reduced to half that of survivors

$$P(M/D) = .5 \times P(M/S) = .5 \times .087 = .0435 \tag{24}$$

and that therefore

$$P(R/D) = 1 - .0435 = .9565 \tag{25}$$

The probability matrix can then be filled out

$$P(R,D) = P(D).\ P(R/D) = .024 \times .9565 = .023 \tag{26}$$

and

$$P(M,D) = P(D).\ P(M/D) = .024 \times .0435 = .001 \tag{27}$$

with P(R) and P(M) being calculated by addition (Figure 12). These probabilities are then used in the simple model proposed here.

In the model we assume the P(R,S), P(M,S), P(R,D) and P(M,D) prospective probabilities remain constant over time, and that to obtain the probabilities of sequences of events stretched over 1,2,3,4 and 5 years we can multiply the probabilities together. That is, we assume constancy and independence of the probabilities. These probabilities are then applied successively over five discrete years to yield probabilities for the event sequences over five years. Some sample probability calculations are

$$P(MS, MS, MS, MS, MS) = .085^5 = .0000044371$$

$$\vdots$$

$$P(MS, MS, RS, MS, RS) = (.085)\,(.085)\,(.891)\,(.085)\,(.891)$$

$$= (.085)^3\,(.891)^2 = .0004875422$$

$$\vdots$$

$$P(MS, MS, RS, MD) \quad = (.085)\,(.085)\,(.891)\,(.001) = .0000064375$$

$$\vdots$$

$$P(RS, RS, RS, RS, RS) = (.891)^5 = .5615501146$$

$$\vdots$$

All the possible sequences and their probabilities are listed in Appendix 1. There are some 94 in all: 32 in which persons survive to the end of the period and 62 in which they die. The survival sequences are generated by a simple binominal process and there must be 2^5 sequences therefore over 5 years. For each survival sequence completed at year 0,1,2,3 and 4 there are two death branches so that the death branches are $2 \times (2^0 + 2^1 + 2^2 + 2^3 + 2^4)$ in number. In general, if there are a outcomes associated with survival, b outcomes associated with death and c time periods involved, the number of permutations or event sequences for survivors will be a^c and for non-survivors will be $b \sum_{x=0}^{c-1} a^x$.

To give the model rather more concrete meaning it has been applied to the estimated population of household heads of Great Britain over the intercensal period 1966-71. The results are displayed as a tree diagram in Figure 13. The estimated 17.8 million heads of

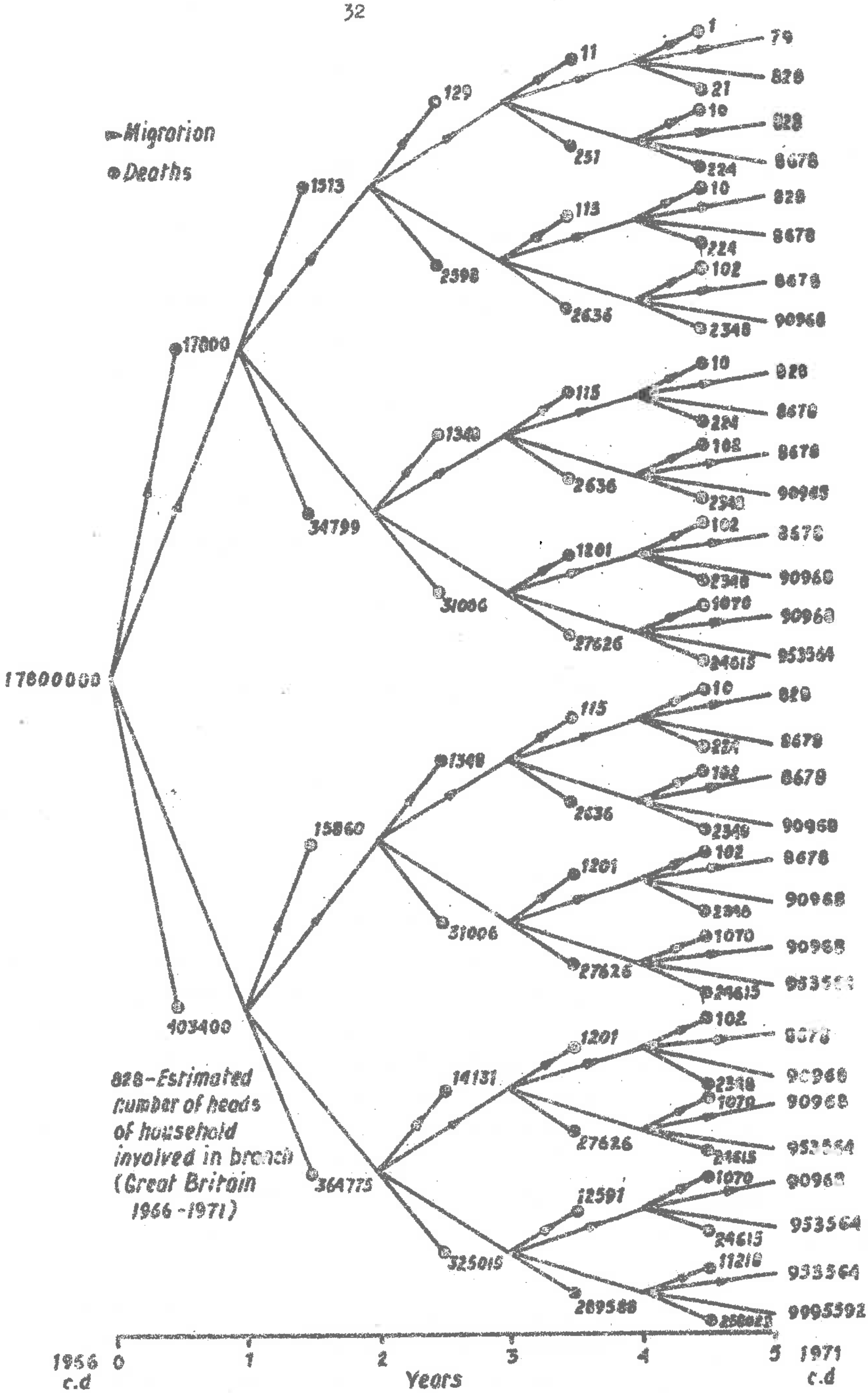


Figure 13 The tree of migration, staying, death and survival states.

household are distributed into the 94 branches of the probability tree. Just under 10 million of them survive without event (the RS,RS,RS,RS,RS sequence); the others suffer a variety of fates.

From the tableau presented in Figure 13 or from the table of probabilities calculated in Appendix 1 one can recalculate the 5 year migrant rate (5 year surviving migrants divided by the surviving population) as a first check on the model (Figure 14). A close approximation is found. One can also form

	Observed in GHS sample	Model
Surviving non-migrants	.646	.634
Surviving migrants	.354	.366

Figure 14 A comparison of observed and model surviving migrant rates

a table (Figure 15) which attempts to assess the relationship between number of migrations and migrants over time. The same relationship of migrations and migrants as observed earlier in Figure 8 is here generated prospectively rather than retrospectively. After five years the population of surviving heads has shrunk to 15.764 millions, of whom 5.768 millions are migrants. These surviving migrants have accomplished at least 6.864 million migrations. Some 2.036 million heads have died, of whom 0.389 million were non-surviving migrants who accomplished 0.431 million migrations. We can divide up the total of migrations into the categories listed earlier in Figure 11. This is done in Figure 16.

		millions	%
I	Required to produce the number of surviving migrants recorded	5.768	79
II	Required to produce the number of non-surviving migrants recorded	0.389	5
III	"Surplus" migrations by surviving migrants	1.096	15
IV	"Surplus" migrations by non-surviving migrants	.042	1
	Total migrations	7.295	100

Figure 16 The breakdown of the migrations generated in the model into components

				Year			Survivors only. Proportions (5)	
Migrant Category	Number of Migrations	1	2	3	4	5	Model	Observed
Surviving migrants	0	.891	.794	.707	.630	.562	.634	.646
	1	.085	.151	.202	.240	.2 8	.302	.234
	2		.007	.019	.034	.051	.058	.067
	3			$.0^{3}61$	.002	.005	.006	.032
	4				$.0^{4}52$	$.0^{3}23$	$.0^{3}26$	.011
	5					$.0^{5}44$	$.0^{3}50$	.011
Non-surviving migrants	0	.023	.043	.062	.078	.093	1.0000	1.0000
	1	.001	.004	.008	.013	.020		
	2		$.0^{4}85$	$.0^{3}40$	.001	.002		
	3			$.0^{5}72$	$.0^{4}41$	$.0^{3}13$		
	4				$.0^{6}62$	$.0^{5}40$		
	5					$.0^{7}60$		
Total		1.000	1.000	1.000	1.000	1.000		
Surv. migrant	Ave.no. of migrations	.087	.174	.261	.348	.435		
Non-Surv. migrants	Ave.no. of migrations	.042	.085	.127	.170	.212		
All migrants	Ave.no.of migrants	.086	.170	.252	.332	.410		

Figure 15 Distribution of migrants by number of migrations made cumulatively over 5 years.

Some 16 percent of the migrations are "surplus" after 5 years, some 15 percent made by surviving migrants and 1 percent by non-surviving migrants. Return migrants do not figure in the table as return migration to the very same home is regarded as a very rare event. Figure 15 also shows the distribution of heads by number of moves observed in the GHS sample of heads. The model appears to have overestimated the number making just one move and underestimated those making more than one move. This stems from the assumption of 1 migration per year per migrant whereas this GHS data indicates that some people must make more moves within a year (an average of 1.274).

7. A multi-regional example

We could regard the model described in section 6 as applicable to a multi-regional situation if the migration probability and staying probability were to be both thought of as migration probabilities in a two region system. In this case half of the migrations made by surviving or non-surviving migrants who made either 2 or 4 migrations would be return migrations (13 percent of the total). However, it would be unrealistic to make this assumption in that situation given that we have placed no bounds on the migrations recorded except that they fall in Great Britain. The likelihood of a migrant returning to the very same home he left must be much smaller than indicated here.

We can, however, apply the section 6 model in a multi-regional context using the data given in Figures 1 and 5 on inter-regional migrants in 1970-71 and 1966-71. We confine attention, to keep things simple, to survivors, and have aggregated the existence-survival portions of the 1970-71 accounts (Figure 5) and 1966-71 (not shown) into a four region by four region table (Figure 17). The transition rate sub-matrices corresponding to these sub-accounts can then be worked out (Figure 18).

The multi-regional version of our section 6 model is then as follows. We assume that the one year transition rates of 1970-71 apply throughout the 1966-71 period and that they are independent. The five year transition rates matrix, $\underline{H}_{ES}(5)$ is then the following sample function of the one year transition rates matrix, $\underline{H}_{ES}(1)$

$$\underline{\underline{H}}_{ES}(5) = \underline{\underline{H}}_{ES}(1)^{\ 5} \qquad (28)$$

This is a simple, first order Markov Chain model. Figure 19 shows our $\underline{\underline{H}}_{ES}(1)$ raised to the fifth power and records the differences between the observed and model rates. The fit of model and observations is not bad, especially compared with those in Figure 10. The differences are, however, systematic: observed diagonal elements are greater than those in the model and off-diagonal elements smaller. This indicates that the average $\underline{\underline{H}}_{ES}(1)$ matrix that would generate the $\underline{\underline{H}}_{ES}(5)$ for 1966-71 would have lower migration rate terms and higher staying terms. If, for example, we compare an earlier set of transition rates for 1965-66 with those of 1970-71 we see that the former contain less migratory behaviour than the latter, with the exception of the Celtic Fringe where the probability of staying and surviving has shifted upwards are the intervening five years (largely the influence of Wales). If we were to compute the average one year matrix that generates the observed five year matrix, that is $(\underline{\underline{H}}_{ES}(5))^{1/5}$, this would probably lie in between the two one year transition rates sub-matrices observed.

One final piece of information can be gleaned from the model. We can make an estimate of the amount of return migration by comparing the fifth powers of the staying and survival elements of the one-year submatrix ($\underline{\underline{H}}_{ES}(1)$ for 1970-71) with the diagonal elements of the $(\underline{\underline{H}}_{ES}(1))^5$ sub matrix:

	$(h^{\varepsilon(i)\sigma\binom{i}{1}})^5$	$(h^{\varepsilon(i)\sigma\binom{i}{5}})$	return migrants
North	.8605	.8609	.0004
Midlands	.8495	.8499	.0004
South	.8517	.8566	.0049
Celtic Fringe	.9390	.9393	.0003

We see that return migrants are a small proportion of the regional populations except in the South where return migrants are half a percent of the surviving stayers.

8. Conclusions

We could take the analysis introduced in sections 6 and 7 a good deal further. We hope, however, that we have been able to show in this paper that the problems of migration measurement are amenable to conceptual analysis and that fairly simple probability models can take us quite far towards an understanding of the processes at work in migration.

1970-71 Aggregated existence-survival accounts

Existence at c.d. 1970 in:	Survival at census date 1971 in: North	Midlands	South	Celtic Fringe	Populations (c.d.1970)
North	14418158	42650	93750	27740	14858305
Midlands	40370	8190609	73080	15120	8461865
South	69620	67740	21838228	42020	22528551
Celtic Fringe	27450	17610	52630	7858075	7957315

1966-71 Aggregated existence survival accounts

Existence at c.d. 1966 in:	Survival at census date 1971 in: North	Midlands	South	Celtic Fringe	Populations (c.d.1966)
North	13049045	135570	275320	85300	14752874
Midlands	118430	7263495	218480	45390	8290708
South	195830	189590	19649230	114920	22175932
Celtic Fringe	78340	53160	149940	6971175	7895805

Notes: North = North, Yorkshire and Humberside and North West;
Midlands = East Midlands, West Midlands;
South = East Anglia, South East and South West;
Celtic Fringe = Wales and Scotland

Figure 17 Aggregated population accounts (existence -survival quadrant). Great Britain, 1970-71 and 1966-71

1970-71 transition rates sub-matrix ($H_{ES}(1)$)

$$\begin{array}{l} \\ \text{North} \\ \text{Midlands} \\ \text{South} \\ \text{Celtic Fringe} \end{array} \begin{array}{c} \begin{array}{cccc} \text{North} & \text{Midlands} & \text{South} & \text{Celtic Fringe} \end{array} \\ \begin{bmatrix} .9704 & .0029 & .0063 & .0019 \\ .0048 & .9679 & .0086 & .0018 \\ .0031 & .0030 & .9694 & .0019 \\ .0034 & .0022 & .0066 & .9875 \end{bmatrix} \end{array}$$

1966-71 transition rates sub-matrix ($H_{ES}(5)$)

$$\begin{array}{l} \\ \text{North} \\ \text{Midlands} \\ \text{South} \\ \text{Celtic Fringe} \end{array} \begin{array}{c} \begin{array}{cccc} \text{North} & \text{Midlands} & \text{South} & \text{Celtic Fringe} \end{array} \\ \begin{bmatrix} .8845 & .0092 & .0187 & .0058 \\ .0143 & .8761 & .0264 & .0055 \\ .0088 & .0085 & .8861 & .0052 \\ .0099 & .0067 & .0190 & .8829 \end{bmatrix} \end{array}$$

Figure 18 The transition rate sub-matrices corresponding to the Figure 17 sub-accounts

$(\underline{H}_{ES}(1))^5$: the fifth power of the 1970-71 transition rates sub-matrix

	North	Midlands	South	Celtic Fringe
North	.8609	.0130	.0282	.0089
Midlands	.0215	.8499	.0383	.0085
South	.0139	.0133	.8566	.0088
Celtic Fringe	.0159	.0103	.0306	.9393

The differences between model and observed, $\underline{H}_{ES}(5)$ rates

	North	Midlands	South	Celtic Fringe
North	.0236	-.0038	-.0095	-.0031
Midlands	.0072	.0262	-.0119	-.0030
South	-.0051	-.0048	.0295	-.0036
Celtic Fringe	-.0060	-.0036	-.0116	-.0564

1965-66 transition rates sub-matrix $(\underline{H}_{ES}(1))$

	North	Midlands	South	Celtic Fringe
North	.9720	.0028	.0048	.0016
Midlands	.0047	.9694	.0071	.0015
South	.0030	.0026	.9696	.0016
Celtic Fringe	.0032	.0024	.0054	.9688

Figure 19 Model transition rates, differences and the 1965-66 transition rates

Appendix 1. The event sequences and their probabilities

	Year 1	Year 2	Year 3	Year 4	Year 5	Probability	Evaluation
1.	MS	MS	MS	MS	MS	$(.085)^5$	.0000044371
2.	MS	MS	MS	MS	RS	$(.085)^4(.891)$	.0000465108
3.	MS	MS	MS	MS	MD	$(.085)^4(.001)$	.0000000522
4.	MS	MS	MS	MS	RD	$(.085)^4(.023)$	.0000012006
5.	MS	MS	MS	RS	MS	$(.085)^4(.891)$	.0000465108
6.	MS	MS	MS	RS	RS	$(.085)^3(.891)^2$	.0004875422
7.	MS	MS	MS	RS	MD	$(.085)^3(.891)(.001)$	.0000005472
8.	MS	MS	MS	RS	RD	$(.085)^3(.891)(.023)$	.0000125853
9.	MS	MS	MS	MD	–	$(.085)^3(.001)$	.0000006141
10.	MS	MS	MS	RD	–	$(.085)^3(.023)$	.0000141249
11.	MS	MS	RS	MS	MS	$(.085)^4(.891)$	.0000465108
12.	MS	MS	RS	MS	RS	$(.085)^3(.891)^2$	.0004875422
13.	MS	MS	RS	MS	MD	$(.085)^3(.891)(.001)$	.0000005472
14.	MS	MS	RS	MS	RD	$(.085)^3(.891)(.023)$	.0000125853
15.	MS	MS	RS	RS	MS	$(.085)^3(.891)^2$	.0004875422
16.	MS	MS	RS	RS	RS	$(.085)^2(.891)^3$	.0051105891
17.	MS	MS	RS	RS	MD	$(.085)^2(.891)^2(.001)$	.0000057358
18.	MS	MS	RS	RS	RD	$(.085)^2(.891)^2(.023)$	.0001319232
19.	MS	MS	RS	MD	–	$(.085)^2(.891)(.001)$	.0000064375
20.	MS	MS	RS	RD	–	$(.085)^2(.891)(.023)$	.0001480619
21.	MS	MS	MD	–	–	$(.085)^2(.001)$	.000007225
22.	MS	MS	RD	–	–	$(.085)^2(.023)$	.000166175
23.	MS	RS	MS	MS	MS	$(.085)^4(.891)$	.0000465108
24.	MS	RS	MS	MS	RS	$(.085)^3(.891)^2$	.0004875422
25.	MS	RS	MS	MS	MD	$(.085)^3(.891)(.001)$	.0000005472
26.	MS	RS	MS	MS	RD	$(.085)^3(.891)(.023)$	.0000125853
27.	MS	RS	MS	RS	MS	$(.085)^3(.891)^2$	.0004875422
28.	MS	RS	MS	RS	RS	$(.085)^2(.891)^3$	.0051105891
29.	MS	RS	MS	RS	MD	$(.085)^2(.891)^2(.001)$	.0000057358
30.	MS	RS	MS	RS	RD	$(.085)^2(.891)^2(.023)$	.0001319232
31.	MS	RS	MS	MD	–	$(.085)^2(.891)(.001)$	.0000064375
32.	MS	RS	MS	RD	–	$(.085)^2(.891)(.023)$	.0001480619
33.	MS	RS	RS	MS	MS	$(.085)^3(.891)^2$	.0004875422
34.	MS	RS	RS	MS	RS	$(.085)^2(.891)^3$	.0051105891
35.	MS	RS	RS	MS	MD	$(.085)^2(.891)^2(.001)$	.0000057358
36.	MS	RS	RS	MS	RD	$(.085)^2(.891)^2(.023)$	.0001319232
37.	MS	RS	RS	RS	MS	$(.085)^2(.891)^3$	.0051105891
38.	MS	RS	RS	RS	RS	$(.085)(.891)^4$	.0535709936
39.	MS	RS	RS	RS	MD	$(.085)(.891)^3(.001)$	.0000601246
40.	MS	RS	RS	RS	RD	$(.085)(.891)^3(.023)$	.0013828653
41.	MS	RS	RS	MD	–	$(.085)(.891)^2(.001)$	.0000674799
42.	MS	RS	RS	RD	–	$(.085)(.891)^2(.023)$	.0015520374
43.	MS	RS	MD	–	–	$(.085)(.891)(.001)$	.000075735
44.	MS	RS	RD	–	–	$(.085)(.891)(.023)$	.001741905
45.	MS	MD	–	–	–	$(.085)(.001)$	.000085000
46.	MS	RD	–	–	–	$(.085)(.023)$	.001955000
47.	RS	MS	MS	MS	MS	$(.891)(.085)^4$	.0000465108
48.	RS	MS	MS	MS	RS	$(.891)^2(.085)^3$	.0004875422
49.	RS	MS	MS	MS	MD	$(.891)(.085)^3(.001)$	.0000005472
50.	RS	MS	MS	MS	RD	$(.891)(.085)^3(.023)$	.0000125853
51.	RS	MS	MS	RS	MS	$(.891)^2(.085)^3$	.0004875422
52.	RS	MS	MS	RS	RS	$(.891)^3(.085)^2$	.0051105891
53.	RS	MS	MS	RS	MD	$(.891)^2(.085)^2(.001)$	.0000057358
54.	RS	MS	MS	RS	RD	$(.891)^2(.085)^2(.023)$	.0001319232
55.	RS	MS	MS	MD	–	$(.891)(.085)^2(.001)$	.0000064375
56.	RS	MS	MS	RD	–	$(.891)(.085)^2(.023)$	.0001480619
57.	RS	MS	RS	MS	MS	$(.891)^2(.085)^3$	.0004875422
58.	RS	MS	RS	MS	RS	$(.891)^3(.085)^2$	.0051105891
59.	RS	MS	RS	MS	MD	$(.891)^2(.085)^2(.001)$	.0000057358

	Year 1	Year 2	Year 3	Year 4	Year 5	Probability	Evaluation
60.	RS	MS	RS	MS	RD	$(.891)^2(.085)^2(.023)$	.0001319232
61.	RS	MS	RS	RS	MS	$(.891)^3(.085)^2$	.0051105891
62.	RS	MS	RS	RS	RS	$(.891)^4(.085)$	.0535709986
63.	RS	MS	RS	RS	MD	$(.891)^3(.085)\ (.001)$	.0000601246
64.	RS	MS	RS	RS	RD	$(.891)^3(.085)\ (.023)$	.0013828653
65.	RS	MS	RS	MD	–	$(.891)^2(.085)\ (.001)$	.0000674799
66.	RS	MS	RS	RD	–	$(.891)^2(.085)\ (.023)$	.0015520374
67.	RS	MS	MD	–	–	$(.891)\ (.085)\ (.001)$	.0000757350
68.	RS	MS	RD	–	–	$(.891)\ (.085)\ (.023$	.0017419050
69.	RS	RS	MS	MS	MS	$(.891)^2(.085)^3$	.0004875422
70.	RS	RS	MS	MS	RS	$(.891)^3(.085)^2$	.0051105891
71.	RS	RS	MS	MS	MD	$(.891)^2(.085)^2(.001)$	.0000057358
72.	RS	RS	MS	MS	RD	$(.891)^2(.085)^2(.023)$	.0001319232
73.	RS	RS	MS	RS	MS	$(.891)^3(.085)^2$	.0051105891
74.	RS	RS	MS	RS	RS	$(.891)^4(.085)$	.0535709986
75.	RS	RS	MS	RS	MD	$(.891)^3(.085)\ (.001)$	.0000601246
76.	RS	RS	MS	RS	RD	$(.891)^3(.085)\ (.023)$	.0013828653
77.	RS	RS	MS	MD	–	$(.891)^2(.085)\ (.001)$	.0000674799
78.	RS	RS	MS	RD	–	$(.891)^2(.085)\ (.023)$	.0015520374
79.	RS	RS	RS	MS	MS	$(.891)^3(.085)^2$	.0051105891
80.	RS	RS	RS	MS	RS	$(.891)^4(.085)$	.0535709986
81.	RS	RS	RS	MS	MD	$(.891)^3(.085)\ (.001)$	.0000601246
82.	RS	RS	RS	MS	RD	$(.891)^3(.085)\ (.023)$	.0013828653
83.	RS	RS	RS	RS	MS	$(.891)^4(.085)$	.0535709986
84.	RS	RS	RS	RS	RS	$(.891)^5$	.5615501146
85.	RS	RS	RS	RS	MD	$(.891)^4(.001)$	.000630247
86.	RS	RS	RS	RS	RD	$(.891)^4(.023)$	.014495682
87.	RS	RS	RS	MD	–	$(.891)^3(.001)$	.000707348
88.	RS	RS	RS	RD	–	$(.891)^3(.023)$	.0162690033
89.	RS	RS	MD	–	–	$(.891)^2(.001)$	.000793881
90.	RS	RS	RD	–	–	$(.891)^2(.023)$	.018259263
91.	RS	MD	–	–	–	$(.891)\ (.001)$	.000891000
92.	RS	RD	–	–	–	$(.891)\ (.023)$	.020493000
93.	MD	–	–	–	–	$(.001)$	.001000000
94.	RD	–	–	–	–	$(.023)$	.023000000

References

Courgeau, D. (1973) Migrants et migrations, Population, 28, pp. 95-129.

Friedlander, D. and Roshier, R.J. (1966) A study of internal migration in England and Wales: Part I, Population Studies, 19, pp. 239-279.

Harris, A.I. and Clausen, R.J. (1967) Labour mobility in Great Britain, 1958-1963, H.M.S.O., SS333, London.

Hyman, G. and Gleave, D. (1976) A reasonable theory of migration, Research Paper 22, Centre for Environmental Studies, London.

Jenkins, J. (1976) Problems of assembling international migration data from census and other sources for use in migration models, population accounts and population forecasts: a preliminary exploration, Working Paper 161, School of Geography, University of Leeds.

Lee, E.S. (1966) A theory of migration, Demography, 3, pp. 47-57. Reprinted in Jackson, J.A. (ed.)(1969) Migration, Cambridge University Press, Cambridge.

Morrison, P.A. and Relles, D.A. (1974) A method for estimating and projecting migration flows for U.S. cities. Paper presented at the British Section Meetings, Regional Science Association, London, 22-23 August, 1974.

Office of Population Censuses and Surveys (O.P.C.S.)(annual) The Registrar General's Statistical Review for England and Wales for the Year .... Part II. Tables, Population, H.M.S.O., London.

Office of Population Censuses and Surveys (O.P.C.S.)(1972) Population projections No. 2, 1971-2011, H.M.S.O., London.

Office of Population Censuses and Surveys (O.P.C.S.) Social Survey Division (1973) The General Household Survey: introductory report, H.M.S.O., London.

Office of Population Censuses and Surveys (O.P.C.S.)(1974-75) Census 1971. England and Wales. Migration Regional Reports (10% sample), H.M.S.O., London.

Office of Population Censuses and Surveys (O.P.C.S.) Edinburgh (1975) Census 1971. Scotland. Migration Tables. Part I (10% sample), H.M.S.O., Edinburgh.

Plessis-Fraissard, M. (1975) Age, length of residence and probability of migration, Working Paper 107, School of Geography, University of Leeds.

Ravenstein, E.G. (1885) The laws of migration, Journal of the Royal Statistical Society, 48, pp. 167-227. Also reprint No. 5-482 in the "Bobbs-Merrill Series in the Social Sciences".

Ravenstein, E.G. (1889) The laws of migration, Journal of the Royal Statistical Society, 52, pp. 241-301. Also reprint No. 5-483 in the "Bobbs Merrill Series in the Social Sciences".

Rees, P.H. (1974). A family of demographic accounts. Working Paper 68, School of Geography, University of Leeds.

Rees, P.H. (1975). A family of demographic accounts and models, revised. Working Paper 119, School of Geography, University of Leeds.

Rees, P.H. (1976). Modelling the regional system: the population component. Working Paper 148, School of Geography, University of Leeds.

Rees, P.H. and Wilson, A.G. (1973). Accounts and models for spatial demographic analysis 1: aggregate population. Environment and Planning, 5, pp. 61-90.

Rees, P.H. and Wilson, A.G. (1976). Spatial population analysis. Edward Arnold, forthcoming.

Rogers, A. (1968). Matrix analysis of inter-regional population growth and distribution. University of California Press, Berkeley and Los Angeles.

Rogers, A. (1975). Introduction to multiregional mathematical demography. John Wiley, London.

Rogers, A. and V.n Rabenau, B (1971). Estimation of inter-regional migration streams from place of birth by residence data. Demography, 8, pp. 185-194.