Working Paper 383

DOES IT REALLY MATTER WHICH MIGRATION DATA YOU USE IN A POPULATION MODEL?

Philip Rees

Paper prepared for the British-Dutch Symposium Publication on Migration, Edited Paul White and Bert Van Der Knaap, Geo Books, Norwich.

School of Geography
University of Leeds
Leeds
LS2 9JT

February 1984

Table of CONTENTS

LIST OF FIGURES

ABSTRACT

The connections between the direct measures of migration commonly used by national statistical agencies and population models are established via careful analysis of a sample of imaginary migration histories. The direct measures involve counts of migrants, of migrations, and of last migrants. Transition population accounts, movement population accounts and previous residence population accounts are developed which correspond to each of the migration measures and which can be used as the basis of regional population projections.

1. INTRODUCTION*

Let us assume that you have been assigned the task of projecting the population of the regions of a country. You realize that one essential in this task is good information on the migration flows among the set of regions. Does it really matter how you measure these migrations and how you incorporate them in your population model?

This paper sets out the argument that it does matter. The second section of the paper develops three alternative migration measures from sets of imaginary migration histories. Sections three, four and five show that each different measure requires its own special treatment in data handling devices called population accounts, and that these accounts are essential preliminaries to population projection.

The final section of the paper reviews the use of each migration measure in a sample of countries and hence the applicability of the three accounting and population projection models.

---

*This paper is a much modified version of the paper the author presented at the Soesterberg conference ("Choices in the construction of regional population projections" which appears in full form in Rees (1984)). At the suggestion of the editors it develops issues covered in the first part of the conference paper in greater detail for a wider audience.

2. THE DIFFERENT MEASURES OF MIGRATION

2.1 A set of imaginary migration histories

Borrowing from ideas set out in Chapter II of Courgeau (1980), let us conceive of a system of two regions, labelled i and j between which persons can make up to three moves in a given time interval (say five years). Region i contains 135,000 people and region j 270,000. What kind of migration histories are possible?

Some eight types of migration history can occur and are labelled A through H in Figure 1. In this diagram each type of migration history is represented by a list of location labels recording where a person is at an instant in time. The diagram records the label only at the start and end of the period, and when things change. The ellipsis ... means that the location continues as previously specified.

(i) Group A make no moves during the time interval. They remain in region i throughout.

(ii) Group B make one move during the interval from region i to region j and then stay put there.

(iii) Group C make two moves: the first from region i to region j, and the second back from region j to region i.

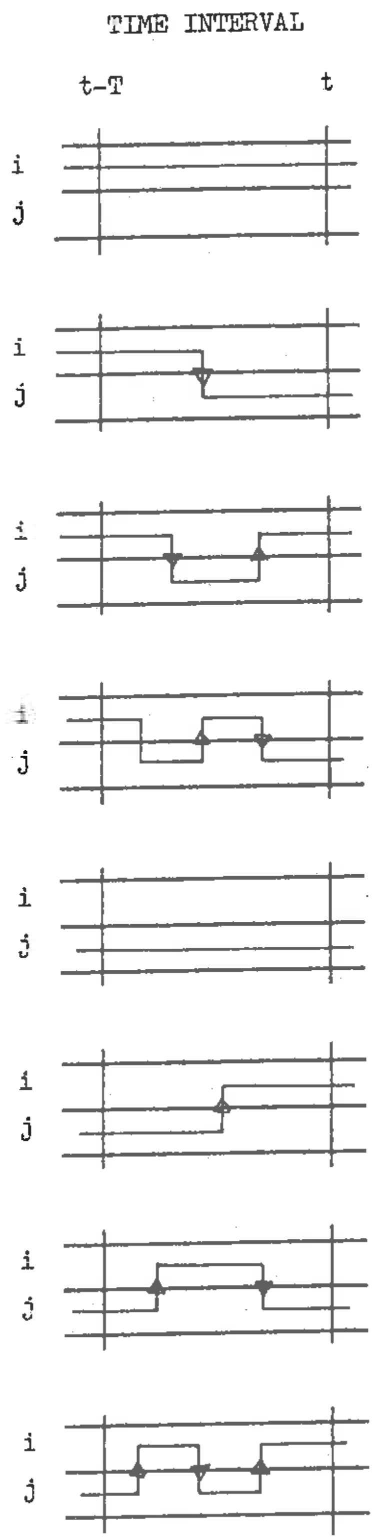
(iv) Group D make three moves: from region i to region j, back again, and then again from region i to region j.

Groups E to H have equivalent migration histories to groups A to D but they start in region j. In a realistic system there would be more regions and more moves, but the principles would not change.

An alternative way of representing these sets of migration histories is set out in Figure 2. To this diagram have been added a set of additional counts. These counts correspond to the three methods in common use for measuring migration.

type of migration history	time at start of interval	time of move one	time of move two	time of move three	time at end of interval	number of histories ('000s)
A	i, ...				..., i	100
B	i, ...,	i,j, ...			..., j	20
C	i, ...,	i,j, ...,	j,i, ...		..., i	10
D	i, ...,	i,j, ...,	j,i,	..., i,j,	..., j	5
E	j, ...				..., j	200
F	j, ...,	j,i, ...			..., i	40
G	j, ...,	j,i, ...,	i,j, ...		..., j	20
H	j, ...,	j,i, ...,	i,j,	..., j,i,	..., i	10

Figure 1. A set of imaginary migration histories



TYPE OF MIGRATION HISTORY	NUMBER OF MIGRATION HISTORIES	MIGRANTS COUNT	MIGRATIONS COUNT	LAST MIGRANT MIGRATION COUNT
A	100	0	0	0
B	20	20	20	20
C	10	0	20	10
D	5	5	15	5
E	200	0	0	0
F	40	40	40	40
G	20	0	40	20
H	10	10	30	10
TOTALS	405	75	165	105

Figure 2. Time-space diagrams for a set of imaginary migration histories

2.2 Methods of measuring migration

The first of these methods has been labelled the migrants (Courgeau 1980) or transitions (Ledent 1980) method. This is based on retrospective questions in end-of-period censuses or surveys of the form "where were you residing T years ago". Migration is measured by comparing the location at time t-T with that at time T.

The second method involves measuring migrations (Courgeau 1980) or movements (Ledent 1980), that is, the event of migrating from one place to another. Migrations or movements counts derive from the compulsory deposit by movers of change of address forms with a national registration system.

The third method, like the first, involves a question asked in a census or retrospective survey but the question involves asking "if you moved during the previous T years, where did you last live?" Courgeau (1980) calls this the last migrants or last migrations method.

Figure 2 shows how we would assemble migration statistics from our imaginary migration histories. The reader can easily check the way the counting is done. According to the first method the volume of migration (1000 is 75; according to the second method it is 165; and according to the third method 105. Note that the counts only agree when either no migrations occur or only one migration occurs during the time interval. The migrants and last migrants counts agree when there are an odd number of migrations made but differ when an even number are made. Neither the migrants count nor the last migrants count exceed the number of migration histories involved in each group, and hence they are counts of persons with particular migration sequences during the time interval. The migrations count can, however, amply exceed the number of migration histories involved: this is because events rather than persons are being counted. The general relationship between the measures is that

migrations between regions i and j	$\geqslant$	last migrations between regions i and j	$\geqslant$	migrants between regions i and j

2.3 Tables of migration flows

How might we assemble these three different counts of migration into tables and link them to counts of population stocks at the start and finish of the time intervals being considered? For the moment we neglect the contribution of births and deaths to population change.

The migration flow tables corresponding to the three measures are set out in Figure 3. The left hand side tables show how we have classified each type of migration history, and the right hand side tables show the corresponding numbers when the numbers in the types in each cell are added together. In the migrations table the types are listed twice when they involve two moves corresponding to the cell row and column labels. No entries are given in the migrations table in the diagonals because a migration is defined as a move that crosses the boundary between region i and region j.

The column totals in each left hand side table are labelled the populations recorded in region i and j in an end of period census, and the same numbers are entered into each right hand side table. For the sake of argument we assume that no census was taken T years before so that the origin totals have to be determined from the other information in the table.

In the migrants table the origin totals are simply the sum of the elements in the table rows. They refer to the populations in region i and j at time t-T, the start of the time interval, under our assumption of a closed population.

To obtain row totals in the migrations table we first compute the diagonal cell entries as the difference between the column total and the

Figure 3. Migration flow tables corresponding to the three migration measures

MIGRANTS TABLE (TRANSITIONS)

FROM \ TO		STATE AT TIME t: REGION i	STATE AT TIME t: REGION j	ORIGIN TOTALS
STATE AT TIME t-T	REGION i	A,C	B,D	
	REGION j	F,H	E,G	
DESTINATION TOTALS		POPN. IN i	POPN. IN j	TOTAL POPN.

← TYPES

NUMBERS ('000s) →

FROM \ TO		STATE AT TIME t: REGION i	STATE AT TIME t: REGION j	ORIGIN TOTALS
STATE AT TIME t-T	REGION i	110	25	135
	REGION j	50	220	270
DESTINATION TOTALS		160	245	405

MIGRATIONS TABLE (MOVEMENTS)

FROM \ TO		STATE AFTER MOVE: REGION i	STATE AFTER MOVE: REGION j	ORIGIN TOTALS
STATE BEFORE MOVE	REGION i	*	B,C,D D,G,H	
	REGION j	C,D,F G,H,H	*	
DESTINATION TOTALS		POPN. in i	POPN. in j	TOTAL POPN.

* not applicable

← TYPES

NUMBERS ('000s) →

FROM \ TO		STATE AFTER MOVE: REGION i	STATE AFTER MOVE: REGION j	ORIGIN TOTALS
STATE BEFORE MOVE	REGION i	65	70	135
	REGION j	95	175	270
DESTINATION TOTALS		160	245	405

LAST MIGRANTS TABLE (LAST MIGRATIONS

FROM \ TO		STATE AT TIME t: REGION i	STATE AT TIME t: REGION j	ORIGIN TOTALS
STATE BEFORE LAST MOVE	REGION i	A	B,D,G	
	REGION j	C,F,H	E	
DESTINATION TOTALS		POPN. IN i	POPN. IN j	TOTAL POPN.

← TYPES

NUMBERS ('000s) →

FROM \ TO		STATE AT TIME t: REGION i	STATE AT TIME t: REGION j	ORIGIN TOTALS
STATE BEFORE LAST MOVE	REGION i	100	45	145
	REGION j	60	200	260
DESTINATION TOTALS		160	245	405

number of migrations into the region. Then the row totals can be computed through summation. These row totals agree with the row totals of the migrants table and can be interpreted as the populations of the regions at the start of the time interval.

In the last migrants table the row totals are computed by addition of the row elements. These numbers, however, differ from those in the two previous tables and have a different interpretation. They are the numbers of people by state previous to the one occupied at the end of the time interval, and thus are not stocks at a particular point in time.

The meaning of the off-diagonal elements in each table should be fairly clear but more attention is needed to the meaning of the diagonal elements. In the migrants table the diagonal cells are made up of persons who fail to move during the time interval, that is, non-movers, plus persons who return to the starting region, that is, return migrants. In the last migrants table the diagonals include only non-movers: return migrants are assigned to the appropriate off-diagonal cells. In the migrations table the diagonals have no substantive meaning. They are simply accounting devices used to obtain the origin populations at the start of the time interval. This can be demonstrated by simply assuming that the people involved in our example made twice as many moves in the next time interval. The resulting migrations table, again under the closed population assumption, would contain one negative diagonal. The diagonals of the migrations table can be interpreted in one sense: they represent the subtraction of the numbers of migrations surplus to those required to place each person in the appropriate transition in the migrants table. There are 70-25 = 45 surplus region i to j migrations and 95-50 = 45 surplus j to i migrations, and the diagonal entries in the migrations table are thus the corresponding migrants table entry less the number of surplus migrations 110-45 = 65 and 220-45 = 175. Note that the number of surplus migrations out of a region is always equal to the number of surplus migrations into a region.

2.4 Adding births and deaths to the migration tables

So far we have looked at migration behaviour in a closed population of persons surveyed at a particular point in time and alive at an earlier point T years ago. The UK Census migration tables refer to such a closed population; but there are three other population types ignored in such a retrospective survey. Figure 4 shows how the four types of population in total are defined. Those in the closed population existing at time t-T and at time t are termed "survivors"; those in the population alive at time t-T but who die before time t are called "non-survivors"; those born during the time interval t-T to t and who are alive at time t are labelled "infant survivors", and the new entrants to the population born during the time interval but also dying before the end are called "infant non-survivors". People in each of these other "transition types" can take part in migration.

Let us assume we have been able to follow the migration histories of these other transition types in the same detail as we were able for survivors. The sets of imaginary migration histories corresponding to the infant survivors, non-survivors and infant non-survivors populations are set out in Figure 5. A "b" has been attached at the start of location label sequences to indicate birth, and a "d" has been added at the end to indicate death. In principle, retrospective surveys or censuses incorporating migration questions should be able to provide details of the migration behaviours of infant survivors, but often the necessary tabulations are not produced. It is unlikely that anything short of a comprehensive, linked registration system could provide details about non-survivors or

FINAL STATE / INITIAL STATE	SURVIVAL AT TIME t	DEATH IN TIME INTERVAL	TOTALS
EXISTENCE AT TIME t-T	SURVIVORS	NON-SURVIVORS	INITIAL POPULATION
BIRTH IN TIME INTERVAL	INFANT SURVIVORS	INFANT NON-SURVIVORS	BIRTHS
TOTALS	FINAL POPULATIONS	DEATHS	TOTAL LIVES

Figure 4. A classification of populations in a time interval

Persons born in time interval who survive to the end (infant survivors)

type of migration history	time at birth in interval	time of move one	time of move two	time of move three	time of end of interval	number of migration histories
A	b,i, ...				..., i	1797
B	b,i, ...,	i,j, ...			..., j	153
C	b,i, ...,	i,j, ...,	j,i, ...		..., i	77
D	b,i, ...,	i,j, ...,	j,i, ...,	i,j,	..., j	38
						2065
E	b,j, ...				..., i	3586
F	b,j, ...,	j,i, ...			..., j	305
G	b,j, ...,	j,i, ...,	j,i, ...		..., i	153
H	b,j, ...,	j,i, ...,	j,i, ...,	i,j,	..., j	76
						4120

Persons existing at the start of the time interval who die before the end (non-survivors)

type of migration history	time at start of interval	time of move one	time of move two	time of move three	time at death in interval	number of migration histories
A	i, ...				... i,d	1540
B	i, ...,	i,j, ...			... j,d	131
C	i, ...,	i,j, ...,	j,i, ...		... i,d	66
D	i, ...,	i.j. ...,	j.i. ...,	i,j,	... j,d	33
						1770
E	j, ...				... j,d	2242
F	j, ...,	j,i, ...			... i,d	191
G	j, ...,	j,i, ...,	i,j, ...		... j,d	95
H	j, ...,	j,i, ...,	i,j, ...,	j,i,	... i,d	47
						2575

Persons born in time interval who die before the end (infant non-survivors)

type of migration history	time at birth in interval	time of move one	time of move two	time of move three	time at death in interval	number of migration histories
A	b,i, ...				...,i,d	22
B	b,i, ...,	i,j, ...			...,j,d	1
C	b,i, ...,	i,j, ...,	j,i, ...		...,i,d	1
D	b,i, ...,	i,j, ...,	j,i, ...,	i,j,	...,j,d	1
						25
E	b,j, ...,	j,i, ...			...,j,d	37
F	b,j, ...,	j,i, ...			...,i,d	2
G	b,j, ...,	j,i, ...,	i,j, ...		...,j,d	1
H	b,j, ...,	j,i, ...,	i,j, ...,	j,i,	,,,,i,d	1
						41

Figure 5. Sets of imaginary migration histories for other transition types

infant non-survivors. However, registration systems which count address transfers will record those of people in all of the transition types, although it will be rare to have knowledge about which transition type for a given time interval they fit into.

We can proceed in the same fashion as for survivors to assemble for these other types of population transition equivalent migrants tables, migrations tables, and last migrants tables (Figure 6). The numbers were computed using the same assignment of migration history types to table cell as was used earlier (see Figure 3).

At this stage we have four migration flow tables for each population transition type. How can they be usefully amalgamated into consolidated tables that show how migration effects population change? This question is tackled in the next section of the paper.

MIGRANTS TABLE (TRANSITIONS)

SURVIVING INFANTS

FROM \ TO		STATE AT TIME t: REGION i	STATE AT TIME t: REGION j	ORIGIN TOTALS
STATE AT BIRTH	REGION i	1874	191	2065
	REGION j	381	3739	4120
DESTINATION TOTALS		2255	3930	6185

NON-SURVIVORS

FROM \ TO		STATE AT DEATH: REGION i	STATE AT DEATH: REGION j	ORIGIN TOTALS
STATE AT TIME t-T	REGION i	1606	164	1770
	REGION j	238	2337	2575
DESTINATION TOTALS		1844	2501	4345

NON-SURVIVING INFANTS

FROM \ TO		STATE AT DEATH: REGION i	STATE AT DEATH: REGION j	ORIGIN TOTALS
STATE AT BIRTH	REGION i	23	2	25
	REGION j	3	38	41
DESTINATION TOTALS		26	40	66

MIGRATIONS TABLE (MOVEMENTS)

SURVIVING INFANTS

FROM \ TO		STATE AFTER MOVE: REGION i	STATE AFTER MOVE: REGION j	ORIGIN TOTALS
STATE BEFORE MOVE	REGION i	1530	535	2065
	REGION j	725	3395	4120
DESTINATION TOTALS		2255	3930	6185

NON-SURVIVORS

FROM \ TO		STATE AFTER MOVE: REGION i	STATE AFTER MOVE: REGION j	ORIGIN TOTALS
STATE BEFORE MOVE	REGION i	1365	405	1770
	REGION j	479	2096	2575
DESTINATION TOTALS		1844	2501	4345

NON-SURVIVING INFANTS

FROM \ TO		STATE AFTER MOVE: REGION i	STATE AFTER MOVE: REGION j	ORIGIN TOTALS
STATE BEFORE MOVE	REGION i	19	6	25
	REGION j	7	34	41
DESTINATION TOTALS		26	40	66

LAST MIGRANTS TABLE (LAST MIGRATIONS)

SURVIVING INFANTS

FROM \ TO		STATE AT TIME t: REGION i	STATE AT TIME t: REGION j	ORIGIN TOTALS
STATE BEFORE LAST MOVE	REGION i	1797	344	2141
	REGION j	458	3586	4044
DESTINATION TOTALS		2255	3930	6185

NON-SURVIVORS

FROM \ TO		STATE AT DEATH: REGION i	STATE AT DEATH: REGION j	ORIGIN TOTALS
STATE BEFORE LAST MOVE	REGION i	1540	259	1799
	REGION j	304	2242	2546
DESTINATION TOTALS		1844	2501	4345

NON-SURVIVING INFANTS

FROM \ TO		STATE AT DEATH: REGION i	STATE AT DEATH: REGION j	ORIGIN TOTALS
STATE BEFORE LAST MOVE	REGION i	22	3	25
	REGION j	4	37	41
DESTINATION TOTALS		26	40	66

Figure 6. Migration flow tables for other transition types

3. ASSEMBLING MIGRANT TABLES INTO TRANSITION ACCOUNTS

Assembly of a population accounts table from the constituent migrants tables is a straightforward task. They are simply attached in the appropriate position, and the necessary row and column totals computed and added to the accounts (Figure 7). If regions i and j constitute a closed system, the row and column totals can be interpreted rather more meaningfully. The row totals of the existence at time t-T section of the accounts represent the initial populations of the regions. The row totals of the births section of the accounts are the total numbers of births in the regions. The column totals of the deaths section represent the total deaths in each region. The totals for the survival columns constitute the final populations of the regions.

These interpretations are essential for the practical construction of transition accounts from data normally available. These data consist of the regional populations, births and deaths together with information on surviving migrants and sometimes infant surviving migrants. The task is then to estimate the missing elements in the accounts table and to adjust the elements to achieve consistent figures.

Figure 8 shows the general structure of a transition accounts table. To estimate the non-surviving migrant terms on the right hand side of the table equations of the form

$$K^{eidj} = K^{eisj}(\tfrac{1}{2}d^{j}/(1-\tfrac{1}{4}d^{j})) \qquad (1)$$

are applied in which weighted death rates are applied to the census derived surviving migrant terms, K^{eisj}. The death rates are derived as

$$d^{j} = K^{e.dj}/\hat{K}^{Dj} \qquad (2)$$

and populations at risk are weighted functions of all transition terms

$$\hat{K}^{Dj} = \sum_{ij} w^{eisj} \; K^{eisj} + \sum_{ij} w^{eidj} \; K^{eidj} \qquad (3)$$

INITIAL STATE \ FINAL STATE		SURVIVAL AT TIME t		DEATH IN INTERVAL		ORIGIN TOTALS
		REGION i	REGION j	REGION i	REGION j	
EXISTENCE AT TIME t-T	REGION i	110,000	25,000	1,606	164	136,770
	REGION j	50,000	220,000	238	2,337	272,575
BIRTH IN TIME INTERVAL	REGION i	1,874	191	23	2	2,090
	REGION j	381	3,739	3	38	4,161
DESTINATION TOTALS		162,255	248,930	1,870	2,541	415,596

Figure 7. Transition accounts for the imaginary migration histories

Initial state in period		Final state in period										Totals
		Survival at time t + T					Death in period t to t + T					
		Internal regions				Outside world	Internal regions				Outside world	
		1	2	...	N	R	1	2		N		
Internal regions	1	$K^{e1\,s1}$	$K^{e1\,s2}$	...	$K^{e1\,sN}$	$K^{e1\,sR}$	$K^{e1\,d1}$	$K^{e1\,d2}$	...	$K^{e1\,dN}$	$K^{e1\,dR}$	$K^{e1..}$
	2	$K^{e2\,s1}$	$K^{e2\,s2}$	...	$K^{e2\,sN}$	$K^{e2\,sR}$	$K^{e2\,d1}$	$K^{e2\,d2}$	...	$K^{e2\,dN}$	$K^{e2\,dR}$	$K^{e2..}$
	⋮	⋮	⋮		⋮	⋮	⋮	⋮		⋮	⋮	⋮
	N	$K^{eN\,s1}$	$K^{eN\,s2}$	...	$K^{eN\,ssN}$	$K^{eN\,sR}$	$K^{eN\,d1}$	$K^{eN\,d2}$	...	$K^{eN\,dN}$	$K^{eN\,dR}$	$K^{eN..}$
Outside world	R	$K^{eR\,s1}$	$K^{eR\,s2}$	...	$K^{eR\,sN}$	Ø	$K^{eR\,d1}$	$K^{eR\,d2}$	...	$K^{eR\,dN}$	Ø	$K^{eR..}$
Totals	.	$K^{e.\,s1}$	$K^{e.\,s2}$	...	$K^{e.\,sN}$	$K^{e.\,sR}$	$K^{e.\,d1}$	$K^{e.\,d2}$	...	$K^{e.\,dN}$	$K^{e.\,dR}$	$K^{e...}$

Notes: 1. Infant accounts have the same structure: a b superscript is substituted for e.
2. Notation:
K = count of persons making transition indicated by superscripts.
e = existence at start of period   s = survival at end of period   d = death before end of period
R = rest of world   . = summation over superscript replaced

Figure 8. A transition accounts table for a period-cohort

The diagonal non-survivor terms are estimated as column residuals

$$K^{eidi} = K^{e.di} - \sum_{j \neq i} K^{ejdi} \tag{4}$$

The diagonal survivor terms can then be derived as residuals

$$K^{eisi} = K^{ei..} - \sum_{j \neq i} K^{eisj} - K^{eidi} - \sum_{j \neq i} K^{eidj} \tag{5}$$

and the final populations are the resulting column sums

$$K^{..sj} = K^{ejsj} + \sum_{i \neq j} K^{eisj} + K^{eRsj} \tag{6}$$

The way in which transition accounts have been used in population projection and the theory and practice of transition accounting have been fully described elsewhere (Rees and Wilson, 1977; Rees, 1981) and a computer program is available for their construction for age-sex disaggregated multiregional population systems (Rees, 1981).

4. ASSEMBLING MIGRATIONS TABLES INTO MOVEMENT ACCOUNTS

4.1 Accounts based on the imaginary migration histories

The migrations tables can be assembled into accounts, distinguishing the four types of population transition, in the same way as the migrants tables were assembled (Figure 9a). However, migration registration systems cannot usually link migration events to subsequent states such as death so that we must simplify the movement accounts for use with conventional data sources. This is done in Figure 9b. The initial state classification is retained as this depends solely on ascertaining the date of birth of the person making a migration at the time of registration, a statistic normally collected. Regional deaths were moved from being column tables to being row entries after disaggregation by initial state, again ascertainable from the age classification of mortality. The migrations tables which occupy the left two columns of the accounts (Figure 9b) contain the sum of the migrations figures from the more detailed tables (Figure 9a) with the diagonals worked out afresh as residuals. These accounts accomplish consistent linkages between initial and final regional populations.

4.2 Movement accounts for the real world

To develop movement accounts for realistic regional systems it is necessary to both shrink and expand the framework arrived at in Figure 9b, and to develop a suitable notation for the terms in the table. Firstly, it simplifies matters a great deal if different period-cohort populations are treated separately. Thus, the infant half of the table becomes the first cohort and the existing half of the table is split up into as many period-cohorts as are necessary for the analysis at hand. Secondly, it is helpful to distinguish between an internal set of regions among which internal migrations occur and the external world. Migrations from internal regions

9a. MOVEMENT ACCOUNTS FOR ALL POPULATION TRANSITIONS

INITIAL STATE \ FINAL STATE			SURVIVAL AT TIME t		DEATHS		ORIGIN TOTALS
			STATE AFTER MOVE		STATE AFTER MOVE		
			REGION i	REGION j	REGION i	REGION j	
EXIST-ENCE AT TIME t-T	STATE BEFORE MOVE	REGION i	65,000	70,000	1,365	405	136,770
		REGION j	95,000	175,000	479	2,096	272,575
BIRTH AT TIME INTER-VAL	STATE BEFORE MOVE	REGION i	1,530	535	19	6	2,090
		REGION j	725	3,395	7	34	4,161
DESTINATION TOTALS			162,255	248,930	1,870	2,541	451,596

9b. MOVEMENT ACCOUNTS RELATED TO CONVENTIONAL DATA SOURCES

	STATE BEFORE MOVE \ STATE AFTER MOVE		STATE AFTER MOVE		DEATHS	ORIGIN TOTALS
			REGION i	REGION j		
EXIST-ENCE AT TIME t-T	STATE BEFORE MOVE	REGION i	64,521	70,405	1,844	136,770
		REGION j	95,479	174,595	2,501	272,575
BIRTH IN TIME INTER-VAL	STATE BEFORE MOVE	REGION i	1,523	541	26	2,090
		REGION j	732	3,389	40	4,161
	DESTINATION		162,255	248,930	4,411	415,596

Figure 9. Movement accounts for the imaginary migration histories

to the external world are called emigrations and migrations to the internal regions from the externalworld are termed immigrations.

For our notation we use a series of different letters to represent the various events, in contrast to the approach in transition accounting where one letter was used for all entries and extensive use was made of subscript notation. The variables comprising a movement accounts table are spelled out in Figure 10. To those in the simple two region table have been added emigrations, E^i_a , and immigrations, I^i_a , which ensure that the population system is closed. No entries are made in the outside world row, outside world column cell or in its deaths neighbour as what goes on inside the outside world is not of interest.

4.3 A population projection model based on movement accounts

The movement accounts represented in Figure 10 can be very simply assembled, once component numbers have been estimated. Simple arithmetic (see below) yields the residuals and final populations. How should such information be used for projection purposes?

A fairly simple but versatile model for projection can be constructed using accounting principles.

4.3.1 Step one: rates

The first step is to compute the various rates of migration and mortality. For movement accounts the population at risk is the average population in the time interval and a good estimate of this is half the initial and final populations:

$$PAR^i_a = \tfrac{1}{2}(IP^i_a + FP^i_a) \qquad (7)$$

STATE BEFORE MOVE		STATE AFTER MOVE						
		INTERNAL REGIONS (DESTINATIONS)				OUTSIDE WORLD	DEATH	TOTALS
		1	2	...	N			
INTERNAL	1	R_a^1	M_a^{12}	...	M_a^{1N}	E_a^1	D_a^1	$IP_a^{1\cdot}$
REGIONS	2	M_a^{21}	R_a^2	...	M_a^{2N}	E_a^2	D_a^2	$IP_a^{2\cdot}$
	⋮	⋮	⋮		⋮	⋮	⋮	⋮
(ORIGINS)	N	M_a^{N1}	M_a^{N2}	...	R_a^N	E_a^N	D_a^N	$IP_a^{N\cdot}$
OUTSIDE WORLD		I_a^1	I_a^2	...	I_a^N	0	0	$I_a^{\cdot}$
TOTALS		FP_a^1	FP_a^2	...	FP_a^N	$E_a^{\cdot}$	$D_a^{\cdot}$	T_a

Definitions of variables

All variables refer to cohort a in a time interval

IP_a^i = initial population of region i

FP_a^i = final population of region i

D_a^i = deaths in region i $D_a^{\cdot}$ = total deaths in all internal regions

E_a^i = emigrations from region i $E_a^{\cdot}$ = total emigration from all interval regions

I_a^i = immigrations to region i I_a^i = total immigration to all interval regions

M_a^{ij} = migrations from region i to region j

R_a^i = residual or balancing term for region i

T_a = total of all inflows/outflows from the internal system

Figure 10. A movement accounts table for a period-cohort

Internal migration rates are given by

$$m_a^{ij} = M_a^{ij} / PAR_a^i \qquad \text{for all } i,j \tag{8}$$

and external migration rates as

$$e_a^i = E_a^i / PAR_a^i \qquad \text{for all } i \tag{9}$$

and

$$i_a^i = I_a^i / PAR_a^i \qquad \text{for all } i \tag{10}$$

Death rates are computed by

$$d_a^i = D_a^i / PAR_a^i \qquad \text{for all } i \tag{11}$$

These rates can be used in the next time interval either as they are or they can be extrapolated in some fashion.

4.3.2 Step two: flows

To compute projected accounts for a future period the rate equations are turned into flow equations:

$$M_a^{ij} = m_a^{ij} \; PAR_a^i \qquad \text{for all } i,j$$

$$E_a^i = e_a^i \; PAR_a^i \qquad \text{for all } i$$

$$I_a^i = i_a^i \; PAR_a^i \qquad \text{"}$$

$$D_a^i = d_a^i \; PAR_a^i \qquad \text{"} \tag{12}$$

4.3.3 Step three: residuals

The items in equation set (12) are those needed to compute the balancing elements:

$$R_a^i = IP_a^i - D_a^i - E_a^i - \sum_{j \neq i} M_a^{ji} \tag{13}$$

4.3.4 Step four: final populations

Then the final populations in the period cohort a can be worked out

$$FP_a^i = R_a^i + \sum_{j \neq i} M_a^{ji} + I_a^i \tag{14}$$

4.3.5 Iteration

However, the population at risks cannot be fully worked out when the model is first computed as the final populations are unknown. The model must therefore recompute the population at risk after step four and repeat equations (12), (13) and (14). This iterative sequence continues until convergence is achieved in, for example, population at risk values. The sequence of population at risk equations is therefore

initially

$$PAR_a^i = \tfrac{1}{2}(IP_a^i + 0) \tag{15}$$

then

$$PAR_a^i = \tfrac{1}{2}(IP_a^i + FP_a^i) \tag{16}$$

with the test for convergence being

$$|PAR_a^i(p) - PAR_a^i(p-1)| \overset{?}{<} 0.5 \tag{17}$$

where p and p-1 refer to iteration numbers.

4.3.6 The infant cohort

Persons born during the time interval are forecast after the "existing" cohorts have been projected. Births by age cohort of parent (or mother) can be projected as

$$B_a^i = b_a^i \; PAR_a^i \qquad \text{for all a in the fertile range} \tag{18}$$

where B_a^i are the births in region i to mothers in period cohort a and

b_a^i is the period cohort specific fertility rate. These births are summed over all fertile cohorts (covering the age range 15-19 through 45-49)

$$B_{.}^i = \sum_a B_a^i \tag{19}$$

and these births entered as initial populations in the infant cohort account

$$IP_1^i = B_{.}^i \tag{20}$$

Then the projection model outlined above can be used.

4.3.7 Model flexibility

The sequence of equations in the movement accounts model could be replaced by a non-iterative set of equations (see Rees, 1984a and 1984b, Willekens and Drewe, 1984) but in such a non-iterative model it is difficult to change the rates or adjust the form of the model. This can be accomplished very easily in the framework of the movement accounts model.

Four choices exist for each component flow variable (M_a^{ij}, E_a^i, D_a^i, I_a^i, B_a^i):

(i) the flows of the previous period are used;

(ii) new flows are input;

(iii) rates from the previous period are used; or

(iv) new rates are input.

For example, migration flows could be forecast for time interval t to t+T that follows interval t-T to T as

$$M_a^{ij}(t,t+T) = M_a^{ij}(t-T,t) \tag{21}$$

or as

$$M_a^{ij}(t,t+T) = \hat{M}_a^{ij} \tag{22}$$

where $\hat{M}_a^{ij}$ are migration flows predicted from an exogenous migration model, or as

$$M_a^{ij}(t,t+T) \;=\; m_a^{ij}\ 9t-T,t)\ \ PAR_a^i(t,t+T) \tag{23}$$

or as

$$M_a^{ij}(t,t+T) \;=\; \hat{m}_a^{ij}(t,t+T)\ \ PAR_a^i(t,t+T) \tag{24}$$

where $\hat{m}_a^{ij}(t,t+T)$ are forecast migration rates for the period t to t+T.

The choice for each component can be made independently which means that mixed flow and rate models are possible and rates for different components can be forecast into the future for different lnegths of time. Full details of a computer program implementing the movement accounts model are given in Rees (1984b) and an example of its application in Rees (1984a).

5. ASSEMBLING LAST MIGRANT TABLES INTO PREVIOUS RESIDENCE ACCOUNTS

5.1 Principles

In section 2 tables of last migrants for the four population transition types were assembled (Figures 3 and 6). The problem with such tables from the point of view of an accounting or projection model is that they fail to connect the initial and final states that regional populations occupy in a time interval. What is missing is a specification of the transitions occuring in the population between initial state and the residence occupied prior to the last migration. This state we will call subsequently "previous residence".

In effect, it is necessary to construct three-dimensional accounts for a time interval. Such a three-dimensional array is shown in Figure 11. The array has three faces (or matrices). The front face with dimensions "previous residence state" by "final state" is what we can construct from observed last migrant tables. The top face with dimensions "initial state" by "final state" is the transition accounts matrix which we need for constructing a projection model. The side face contains the missing link - a matrix with dimensions "initial state" and "previous residence state". There are two stages,therefore, in using the last migrant measure to construct population accounts and projection models. In the first stage the available information on last migrants, populations, births and deaths is used to make an estimate of the elements of the three-dimensional array (the cube of Figure 11). In the second stage the appropriate elements of the array are aggregated to yield the "target matrix" face IF.

5.2 Accounts based on the imaginary migration histories

To simplify the problem of assigning each type of migration history to a cell in the three-dimensional last migrant accounts, the migration histories

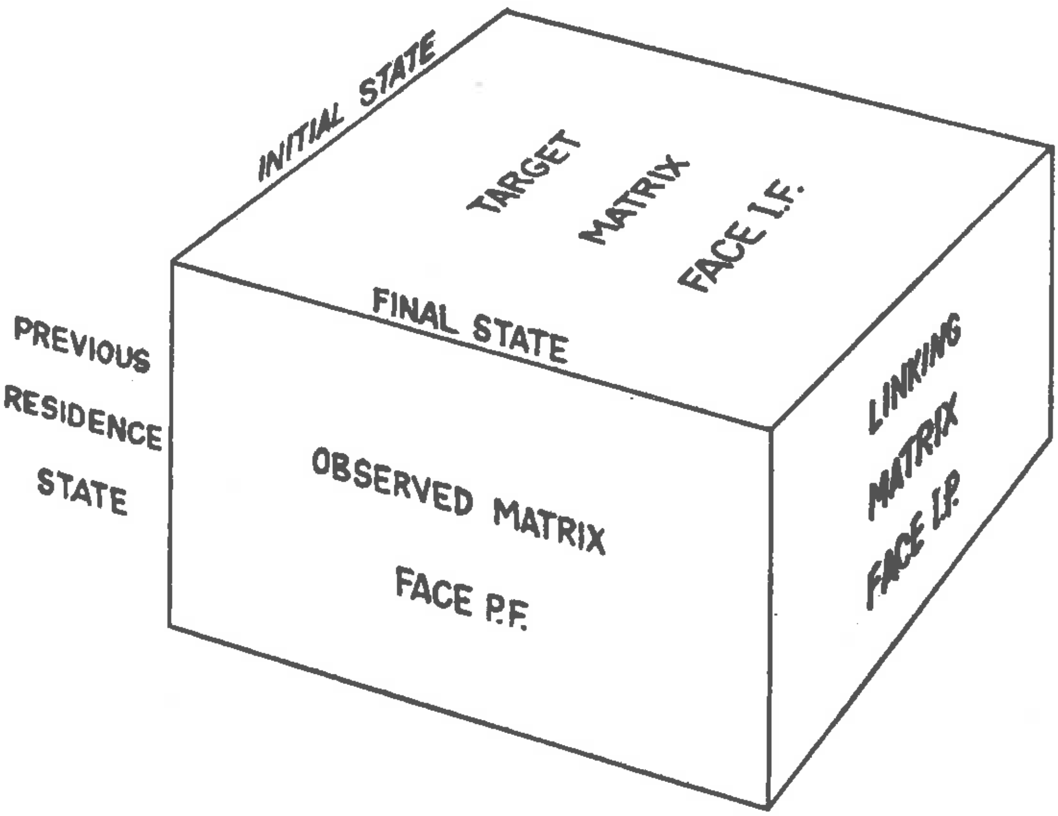


Figure 11. A three dimensional representation of previous residence population accounts

have been reassembled in Figure 12 in initial state - previous residence - final state sequences. Note that persons making one or three moves between regions i and j are assigned to the same sequence. Note also that location sequences of the type "i, j, j" or "j, i, i" are not allowed. If your previous residence location is the same as your end of period location in a time interval, your initial location cannot be different as it would then be classified as your previous residence!

Figure 13 shows the contents of the last migrant array and three associated faces for the imaginary migration histories. The previous residence state has been combined with the initial state for convenience. The disallowed state sequences are represented as empty cells. Some four final states, two previous residence states, and four initial states are recognised. In practice, the states "existence at time t-T" and "birth in time interval" can be attached to any dimension since they depend only on an age classification of last migrants, but it is most convenient to show them as initial states. Thus, face PF represents the aggregation of the four last migrant tables assembled earlier (Figures 3 and 6): survivors and infant survivors are combined; non-survivors and infant non-survivors are combined.

5.3 A sketch of a previous residence accounts model

How should we go about estimating the contents of the last migrant accounts array? There is insufficient space here to spell out the model equations in detail but some of the principles involved can be sketched in.

The accounts variables are laid out in general form in Figure 14. The meaning of the accounts entries is as follows:

INITIAL STATE	PREVIOUS RESIDENCE	FINAL STATE	MOVES	TYPE OF MIGRATION HISTORY	NUMBER OF MIGRATION HISTORIES
EXIST-SURVIVE					
i	i	i	0	A	100,000
i	j	i	2	C	10,000
j	j	i	1,3	F,H	50,000
j	j	j	0	E	200,000
j	i	j	2	G	20,000
i	i	j	1,3	B,D	25,000
EXIST-DIE					
i	i	i,d	0	A	1,540
i	j	i,d	2	C	66
j	j	i,d	1,3	F,H	238
j	j	j,d	0	E	2,242
j	i	j,d	2	G	95
i	i	j,d	1,3	B,D	164
BORN-SURVIVE					
b,i	i	i	0	A	1,797
b,i	j	i	2	C	77
b,j	j	i	1,3	F,H	381
b,j	j	j	0	E	3,586
b,j	i	j	2	G	153
b,i	i	j	1,3	B,D	191
BORN-DIE					
b,i	i	i,d	0	A	22
b,i	j	i,d	2	C	1
b,j	j	i,d	1,3	F,H	3
b,j	j	j,d	0	E	37
b,j	i	j,d	2	G	1
b,i	i	j,d	1,3	B,D	2

Figure 12. Migration histories classified by initial state, previous residence and final state

ARRAY IPF INITIAL STATE BY PREVIOUS RESIDENCE BY FINAL STATE

INITIAL STATE: LIFE STATE	INITIAL STATE: REGION	PREVIOUS RESIDENCE STATE	FINAL STATE: SURVIVAL AT TIME t: REGION i	FINAL STATE: SURVIVAL AT TIME t: REGION j	FINAL STATE: DEATH IN INTERVAL: REGION i	FINAL STATE: DEATH IN INTERVAL: REGION j	ORIGIN TOTALS
EXIST-ENCE AT TIME t-T	REGION i	REGION i	100,000	25,000	1,540	164	126,704
		REGION j	10,000	Ø	66	Ø	10,066
	REGION j	REGION i	Ø	20,000	Ø	95	20,095
		REGION j	50,000	200,000	238	2,242	252,480
BIRTH IN TIME INTER-VAL	REGION i	REGION i	1,797	191	22	2	2,012
		REGION j	77	Ø	1	Ø	78
	REGION j	REGION i	Ø	153	Ø	1	154
		REGION j	381	3,586	3	37	4,007
DESTINATION TOTALS			162,255	248,930	1,870	2,541	415,596

FACE PF

PREVIOUS RESIDENCE	FINAL STATE: SURVIVAL AT TIME t: REGION i	FINAL STATE: SURVIVAL AT TIME t: REGION j	FINAL STATE: DEATH IN INTERVAL: REGION i	FINAL STATE: DEATH IN INTERVAL: REGION j	ORIGIN TOTALS
REGION i	101,797	45,344	1,562	262	148,965
REGION j	60,458	203,586	308	2,279	266,631
DESTINATION TOTALS	162,255	248,930	1,870	2,541	415,596

FACE IP

INITIAL STATE: LIFE STATE	INITIAL STATE: REGION	PREVIOUS RESIDENCE: REGION i	PREVIOUS RESIDENCE: REGION j	ORIGIN TOTALS
EXISTENCE AT TIME t - T	REGION i	126,704	10,066	136,770
	REGION j	20,095	252,480	272,575
BIRTH IN TIME INTERVAL	REGION i	2,012	78	2,090
	REGION j	154	4,007	4,161
DESTINATION TOTALS		148,965	266,631	415,596

FACE IF

INITIAL STATE: LIFE STATE	INITIAL STATE: REGION	FINAL STATE: SURVIVAL AT TIME t: REGION i	FINAL STATE: SURVIVAL AT TIME t: REGION j	FINAL STATE: DEATH IN INTERVAL: REGION i	FINAL STATE: DEATH IN INTERVAL: REGION j	ORIGIN TOTALS
EXISTENCE AT TIME t-T	REGION i	110,000	25,000	1,606	164	136,770
	REGION j	50,000	220,000	238	2,337	272,575
BIRTH IN TIME INTERVAL	REGION i	1,874	191	23	2	2,090
	REGION j	381	3,739	3	38	4,161
DESTINATION TOTALS		162,255	248,930	1,870	2,541	415,596

Figure 13. Previous residence accounts for the imaginary migration histories

INITIAL STATE		PREVIOUS RESIDENCE	SURVIVAL AT END OF INTERVAL					DEATH IN INTERVAL					TOTALS
			REGION 1	REGION 2	...	REGION N	REGION R	REGION 1	REGION 2	...	REGION N	REGION R	
EXISTENCE AT START OF INTERVAL	REGION 1	REGION 1	K^{e1p1s1}	K^{e1p1s2}	...	K^{e1p1sN}	K^{e1p1sR}	K^{e1p1d1}	K^{e1p1d2}	...	K^{e1p1dN}	K^{e1p1dR}	$K^{e1p1..}$
		⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
		REGION N	K^{e1pNs1}	K^{e1pNs2}	...	∅	K^{e1pNsR}	K^{e1pNd1}	K^{e1pNd2}	...	∅	K^{e1pNdR}	$K^{e1pN..}$
		REGION R	K^{e1pRs1}	K^{e1pRs2}	...	K^{e1pRsN}	∅	K^{e1pRd1}	K^{e1pRd2}	...	K^{e1pRdN}	∅	$K^{e1pR..}$
	⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	REGION N	REGION 1	∅	K^{eNp1s2}	...	K^{eNp1sN}	K^{eNp1sR}	∅	K^{eNp1d2}	...	K^{eNp1dN}	K^{eNp1dR}	$K^{eNp1..}$
		⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
		REGION N	K^{eNpNs1}	K^{eNpNs2}	...	K^{eNpNsN}	K^{eNpNsR}	K^{eNpNd1}	K^{eNpNd2}	...	K^{eNpNdN}	K^{eNpNdR}	$K^{eNpN..}$
		REGION R	K^{eNpRs1}	K^{eNpRs2}	...	K^{eNpRsN}	∅	K^{eNpRd1}	K^{eNpRd2}	...	K^{eNpRdN}	∅	$K^{eNpR..}$
	REGION R	REGION 1	∅	K^{eRp1s2}	...	K^{eRp1sN}	K^{eRp1sR}	∅	K^{eRp1d2}	...	K^{eRp1dN}	K^{eRp1dR}	$K^{eRp1..}$
		⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
		REGION N	K^{eRpNs1}	K^{eRpNs2}	...	∅	K^{eRpNsR}	K^{eRpNd1}	K^{eRpNd2}	...	∅	K^{eRpNdR}	$K^{eRpN..}$
		REGION R	K^{eRpRs1}	K^{eRpRs2}	...	K^{eRpRsN}	0	K^{eRpRd1}	K^{eRpRd2}	...	K^{eRpRdN}	0	$K^{eRpR..}$
TOTALS			$K^{e.p.s1}$	$K^{e.p.s2}$	...	$K^{e.p.sN}$	$K^{e.p.sR}$	$K^{e.p.d1}$	$K^{e.p.d2}$	...	$K^{e.p.dN}$	$K^{e.p.dR}$	$K^{e.p...}$

Figure 14. Previous residence accounts for a period-cohort : algebraic variables

K^{eipjsk} = persons in region k at the end of the time interval, in previous residence region j before the last migration and in initial location region i

K^{eipjdk} = persons dying in region k during the time interval, in previous residence region j before the last migration and in initial location region i.

Sequences eipksk and eipkdk are not allowed.

From these variables we can define a rate at which surviving persons are admitted to final states from previous residences from the data provided in last migrant tables.

$$a^{jk} = (K^{e.pjsk} / K^{e.p.sk}) \qquad (25)$$

We can then assume that these rates also apply to the initial state to previous residence transition to yield an estimate of the accounts variables

$$K^{eipjsk} = a^{ij} K^{e.pjsk} \qquad (26)$$

This equation assumes that only two moves take place during the time interval. If three moves are possible (or known to be empirically important) the estimating equation could be extended to

$$K^{eipjsk} = \sum_m a^{im} a^{mj} K^{e.pjsk} \qquad (27)$$

The non-survivor terms could be estimated using equations similar to those involved in the transition accounts (see section 3).

So, in principle, though it would involve a lot of estimating equations, it would be possible to reconstruct from the last migrants tables sets of transition accounts. Courgeau's (1980) dismissal of this type of migration statistic as of little utility thus requires some revision, although much further development of this model is needed.

6. CONCLUSIONS

This chapter has attempted to expose in more depth the meaning of the different direct measures of migration in common use in surveys, censuses and registration systems in very many countries. An awareness of the meaning of each measure is important for researchers using migration data either for descriptive purposes or for studying migratory behaviour or for projecting the population.

To what countries are the remarks in section 2 (complete migration histories), section 3 (transitions), section 4 (movements) and section 5 (last migrants) most relevant? Figure 15 gives a picture of the availability of the different types of migration measure across a sample of countries. The classification is based mainly on a reading of secondary research monographs (such as the Migration and Settlement series of the International Institute for Applied Systems Analysis).

Among this set of mainly developed nations, the commonest measure used is the migrations/movements measure. A smaller set of countries gather mainly transition data via censuses or national surveys. Although only one has been placed in last migrant/migrations class it is probable many more countries use this question in their censuses since the question is recommended by the United Nations.

There are a number of countries which collect more than one type of statistic: Japan collects registration data on migration and also asked census questions on migration in 1960 and in 1970, using the transition measure in the first and last migrant measure in the second (Nanjo, Kawashima and Kuroda, 1982, Appendix E). In the Soviet Union a two year transition question was asked in the 1970 census (Thomas and Rees, 1979) and registration data are also available (Soboleva, 1980). In the United Kingdom a "one year"

Type of migration data	Countries publishing migration data
Complete migration histories (linked registers)	Sweden A
Movement data (Registers)	Sweden B, Finland, GDR, Netherlands Hungary, FRG, Poland, Bulgaria, Czechoslovakia, Soviet Union A, Italy, United Kingdom A, Japan A
Transition data (censuses, surveys)	United Kingdom B, Canada, Austria, United States, Australia, France, Soviet Union B, Japan B
Last migrant data (censuses, surveys)	Japan C

Sources: The classification was deduced from the description of migration data given in secondary sources.

Sweden B, Finland, GDR, Netherlands, Hungary, FRG, Poland, Bulgaria, Czechoslovakia, Soviet Union A, Italy: Rees and Willekens, 1984.
United Kingdom B, Canada, Austria, United States, France: Rees and Willekens, 1984.
Soviet Union B: Thomas and Rees, 1979.
Sweden A: Ledent and Rees, 1984.
Japan, A, B, C: Nanjo, Kawashima and Kuroda, 1982, Appendix E.
United Kingdom A: Rees, 1984; Stillwell, 1984.
Australia: McKay and Whitelaw, 1978.

Figure 15. A classification of national migration data sets

transition question was asked in 1961, 1966, 1971 and 1981, "five year" questions were asked in 1966 and 1971, and from 1974 onwards surrogate movement data is available from the National Health Service Central Register (Stillwell, 1984). Researchers using the migration history data made available by the Swedish and Norwegian statistical offices can, of course, choose whichever of the aggregate measures they please.

So, how do we answer the question posed in the title of this Chapter? For those countries appearing only once in the Figure 15 classification, researchers have no choice as to which migration data to use. What they do have is a choice of accounting and projection model frameworks and they should clearly choose the appropriate one. For those countries where there is a choice of migration data, the choice is not critical as long as the right accounting framework and projection model are matched with the data chosen. Setting aside the (considerable) possibility of variable errors in the different measures, the analysis of this chapter suggests you should obtain the same projection for the same inputs, irrespective of data choice as long as the appropriate model is used. We should therefore turn around the initial question and ask instead "does it really matter which population accounting and projection model you use?". The answer can only be a resounding "yes".

REFERENCES

Courgeau, D. (1980) Analyse quantitative des migrations humaines, Masson, Paris.

Ledent, J. (1980) Multistate life tables: movement versus transition perspectives. Environment and Planning A, 12, 533-562.

Ledent, J. and Rees, P.H. (1984) Life tables. Chapter 10 in Rogers, A. and Willekens, F. (ed) Migration and settlement: a multiregional comparative study. Forthcoming.

McKay, J. and Whitelaw, J.S. (1978) Internal migration and the Australian urban system. Progress in Planning, 10, 1-83.

Nanjo, A., Kawashima, T. and Kuroda, T. (1982) Migration and settlement: 13. Japan. Research Report RR-82-5, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Rees, P.H. (1981) Accounts based models for multiregional population analysis: methods, program and users' manual. Working Paper 295, School of Geography, University of Leeds.

Rees, P.H. (1984a) Choices in the construction of regional population projections. Chapter 6 in Woods, R. and Rees, P. (ed) Developments in spatial demography. Forthcoming.

Rees, P.H. (1984b) Spatial population analysis using movement data and accounting methods: theory, models the MOVE program and examples. Working Paper, School of Geography, University of Leeds. Forthcoming.

Rees, P.H. and Willekens, F. (1984) Accounts and data. Chapter 2 in Rogers, A. and Willekens, F. (ed) Migration and settlement: a multiregional comparative study. Forthcoming.

Rees, P.H. and Wilson, A.G. (1977) Spatial population analysis. Edward Arnold, London.

Soboleva, S. (1980) Migration and settlement: 8. Soviet Union. Research Report RR-80-36, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Stillwell, J.C.H. (1984) Migration between metropolitan and non-metropolitan regions in the UK. Chapter 7 in Woods, R. and Rees, P.H. (ed) Developments in spatial demography. Forthcoming.

Thomas, D. and Rees, P.H. (1980) The analysis of population change in the republics of the Soviet Union using multiregional accounting and projection methods, Part I, 1968-76. Working Paper 276, School of Geography, University of Leeds.

Willekens, F. and Drewe, P. (1984) A multiregional model for regional demographic projection. Chapter 13 in Willekens, F. and Ter Heide, H. (ed) Demographic research and spatial policy. Academic Press. Forthcoming.