Working Paper 378

CHOICES IN THE CONSTRUCTION OF REGIONAL POPULATION PROJECTIONS

Philip Rees

School of Geography
University of Leeds
Leeds LS2 9JT

This version: March 1984

Paper prepared for the Anglo-Dutch Symposium on Migration, Soesterberg, Netherlands, 14-16 September, 1983

CONTENTS

ACKNOWLEDGEMENTS

I would like to thank Frans Willekens of the Netherlands Inter-University Demographic Institute for his kind invitation to visit NIDI and to develop a program to construct movement accounts and project regional populations. We hope to finish the application of this work to the Netherlands inter-provincial system in the near future. My colleague John Stillwell also deserves my grateful thanks for his good offices in obtaining the NHSCR migration data described in the paper, and for his work in turning them into a data set for our mutual 20 zone system.

ABSTRACT

The paper reviews the choices facing a person wishing to project the populations of a set of regions and presents some new alternatives. These options are based on a new accounting framework and model, that for movement data, which are described in detail. The new framework makes possible experimentation for a UK 20 zone system with different ways of handling the accounts and in dealing with the world outside the region set of interest. Given the same view of regional population change in the base period and the same view of the way components of change will develop in the future it is still possible to come to widely differing conclusions about the projected population of regions.

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Overview

Projecting the population of geographical areas into the future is an exercise now routinely carried out at both national and local levels. The projected numbers of people have a profound effect on the demand for public services, as Craig (1983) has recently demonstrated so elegantly using national population projections.

Of course, population projections are always subject to error, and must be continually revised to take account of the most up-to-date information on trends in the components of change that "drive" the projections. But the forecaster needs also to be aware of the likely sources of variation, besides those of real world events, that will affect the projections made. Some of these sources of variation are well known; others are not. The intention of this paper is to explore the likely effects of two sources that involve "model" and "system" design, and which to date have been neglected. These two sources are the accounting framework employed in the population projection, and the way in which the system of regions being studied is closed. A set of empirical experiments using a United Kingdom (UK) twenty zone system are used to tease out the significance of these sources of variation. Section 2 of the paper reviews all sources of variation in population projections.

Section 3 of the paper describes the alternative accounting frameworks that underpin population projections, describing in detail that based on movement data. The fourth section spells out the variety of ways accounts can be built, using the component information in different ways. Section 5 specifies the different ways in which the system may be closed. A brief description of the demographic information system of raw data files, estimation procedures and processed data files that had to be constructed is given in the sixth section. Section 7 outlines the research design governing the projection experiments, the results of which are reviewed and evaluated in section 8.

In the remainder of this introduction some essential preliminary points are made concerning the age-time plan involved in population projection and the crucial distinctions between the different migration measures are set out.

1.2 The age-time plan of population projection

Any regional population projection model must work with the population disaggregated into regular age groups, because most services have client groups concentrated in a particular age span. This means that the most natural model to use will involve "cohort survival". It is essential in projection to focus on cohorts, as this is the way in which a projection model works. Unfortunately, data for most components of change in the UK are tabulated only by age at time of the event.

To understand the importance of the age-time plan used it is essential to think "diagrammatically". Figure 1 sets out four age-time plans in a series of Lexis diagrams. The Lexis diagram labelled "type I" shows the way in which demographic events in the UK are usually recorded, by age group at time of the event, the pecked square areas on the diagram. The "type II" Lexis diagram shows the way in which demographic events should be counted for entry into projection models, that is in the pecked parallelogram areas. This is because persons, in a cohort, move from being in one age group at the start of a time interval to being in the next at the end. Thus, the person with lifeline PQ starts aged 0-4 at time t and moves to be aged 5-9 at time t+5. The third type of Lexis diagram involves cohorts observed over two consecutive periods between two exact ages. The pecked parallelogram areas enclose the relevant events. It is in measuring life expectancy using the life table model that the "type III" age-time plan is relevant. We would be interested in measuring the probability of persons with similar lifelines to the PQR line in the diagram of surviving from age 5 to age 10, for example.

The fourth age-time plan shows how demographic events ideally should be observed: by combinations of age group and cohort - the pecked triangular areas on the diagram. Events in these triangles can be reassembled for projection purposes or for life table purposes, without the necessity for intermediate estimation equations.

A second essential preliminary concept to explain is the distinction

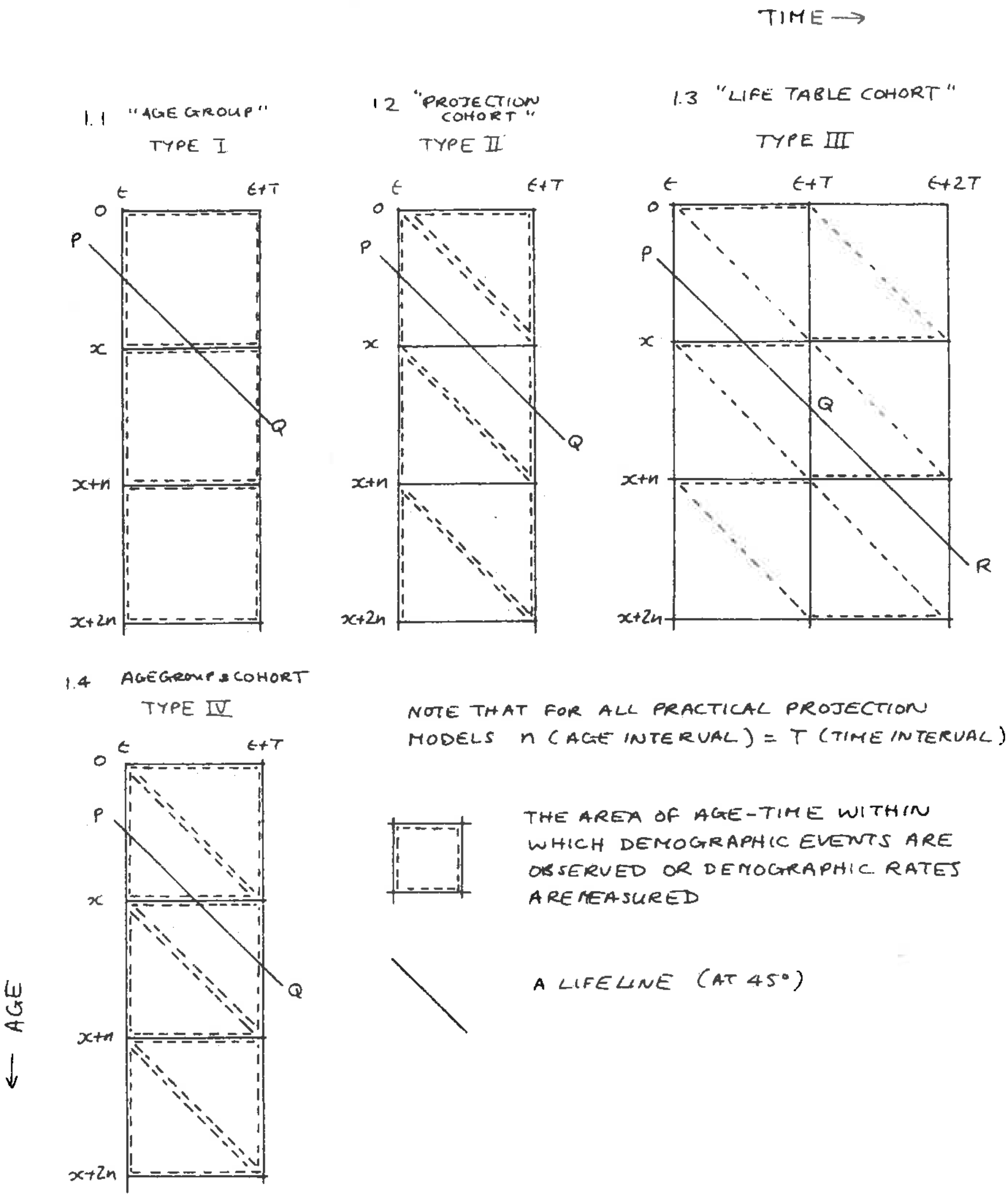


FIGURE 1 ALTERNATIVE AGE-TIME OBSERVATION PLANS

between the different methods of observing migration behaviour. At the top of Figure 2 the migration histories of five individuals are traced as they move between two regions (Figure 2.1). Our migration measurement instruments yield partial extracts of this full picture. A fixed period question in a census yields a picture (Figure 2.2) of transitions retrospectively by comparing current location with location T years ago (persons 1,2) or, if you are lucky, at birth (person 5). One migrant and one infant migrant are counted. Registers count the transfer events or moves - three out of region i, three out of region 1 (Figure 2.3). A question on last residence in a census yields a third measure of migration (Figure 2.4). In this diagram two last migrations are measured from region j to region i and one in the opposite direction. The importance of these three migration concepts is that corresponding to each is a set of multiregional population accounts and a multiregional projection model.

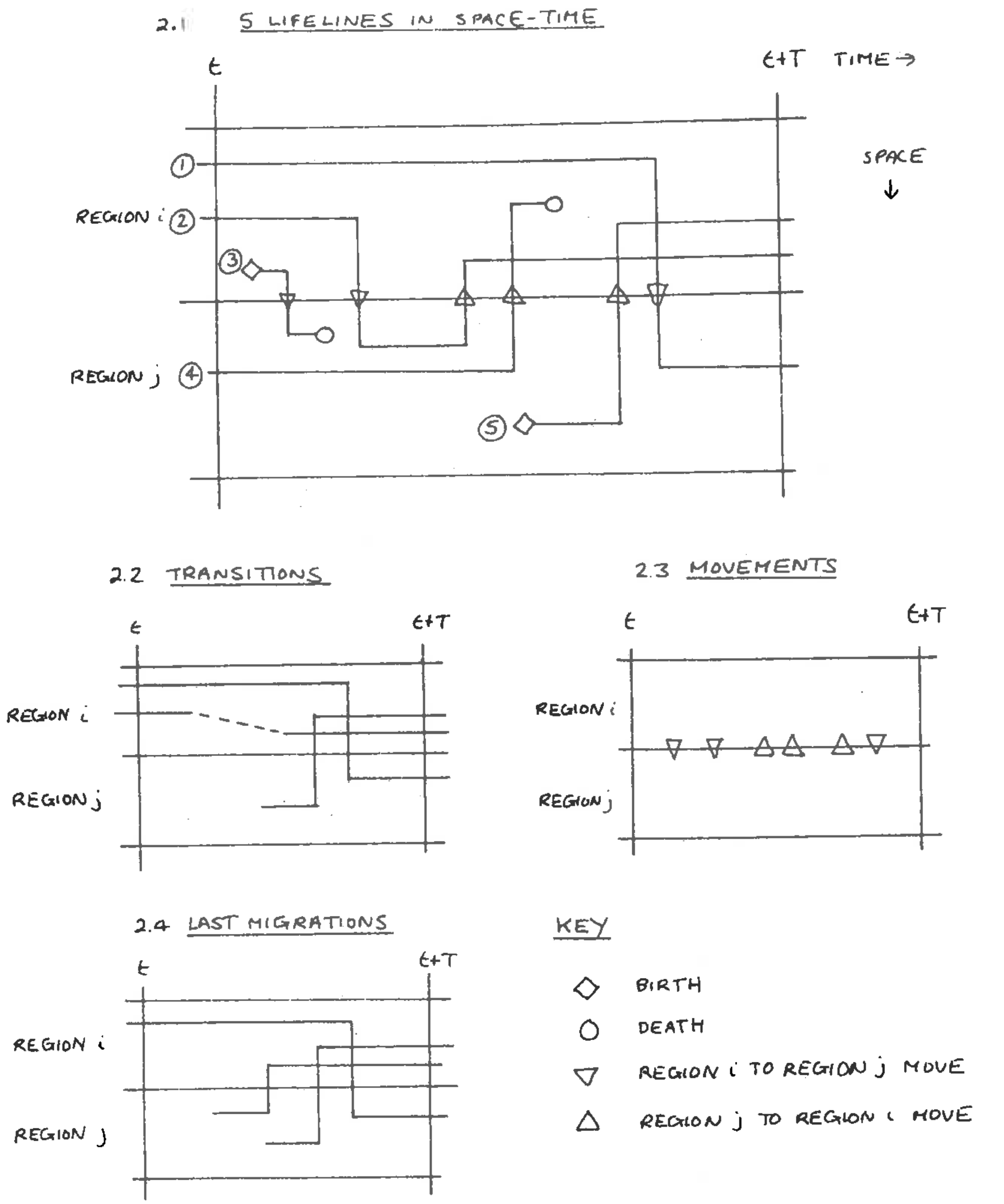


FIGURE 2 THE DIFFERENT KINDS OF MIGRATION MEASURE

2. SOURCES OF VARIATION IN POPULATION PROJECTION

We can identify the following sources of variation:

(i) the assumptions about the way in which the component demographic rates or flows input to the projection are likely to change in the future;

(ii) the influence of events outside the demographic system which disturb its momentum and alter the direction of change;

(iii) the level of aggregation and /or decomposition of the regional system being studied;

(iv) the way in which the regional system under study is closed; and

(v) the way in which the data for the base period are assembled and made consistent.

2.1 Changes in input assumptions

The first influence is obviously of crucial importance and is well recognized. Projections based on alternative fertility levels are regularly carried out (e.g.Central Policy Review Staff, 1976). Occasionally, the effect of curing cancer or of reducing cardiovascular death rates are explored, and although, overall the effect on population size is small (Craig,1983), such changes can affect the elderly age groups a lot. Only a few researchers in the UK have explored the consequences of different mobility levels (Stone, 1970) because only recently has a good time series become available (Ogilvy, 1980; Stillwell, 1983).

2.2 Events outside the demographic system

Events outside the demographic system clearly are very important in influencing component rates. Social norms, ideas of family life and economic prospects are all likely to affect fertility levels, perhaps in a phased inter-generational fashion (Easterlin, 1980). The slowing of economic growth since 1974 has had a profound effect on mobility levels and hence on projected populations. Major changes in interregional and international competitiveness will also affect future migration flows. However, interesting and vital though these issues are, they must be left aside in the present paper, although, with the demographic information established for the experiments described later, it should be possible to begin exploration of some of them.

2.3 Aggregation/decomposition

The third source of variation - aggregation/decomposition - has been explored in a seminal paper by Rogers (1976). Rogers has long argued that regional population projections should be carried out using a multiregional model that incorporates gross migration rates amongst regions, rather than by using a model based on net migration flows or rates (see Rogers, 1983 for the full arguments). This strong and valid case has still to win many converts in national demographic offices, although other branches have recognized the need for a multiregional projection model (Martin and Voorhees Associates and John Bates Services, 1981).

Rogers (1976) explores the differences between the projected results of a full 9 by 9 multiregional model for the USA and a series of 8 alternative models, ranging from a components of growth model (no age disaggregation) through to a model in which the system is decomposed into two sets. Within each set the regions interact but between the sets the flows and the rates are pooled. Table 1 sets out the performance statistics for each model compared with the full multiregional model. The "best buy" among the alternatives is the third model, involving bi-regional aggregation, because of the simplicity of its data requirements: just in- and out-migration to the rest of the country by age for each region. The eighth model, decomposition B with bi-regional aggregation, performs better but demands a full interregional migration matrix for implementation. In such a situation it would be simpler to use a full multiregional model since the data requirements are the same.

2.4 The importance of external migration and the issue of system closure

However, in the Rogers' school of projection attention has been concentrated on intra-national multiregional systems, assumed closed to outside influences. The theoretical influence of external migration on national populations has recently been explored by Alexander (1983). She shows that projection trajectories in a "this country-rest of the world" system need not necessarily exhibit the stable (that is, constant rate) growth or decline of internal regional systems: stationarity can be achieved when natural gain within the country is balanced by net emigration

TABLE 1. Comparisons of Rogers' different aggregations/ decompositions with the full 9 x 9 multiregional model at 2008

no.	Model title	Difference in US total population (1000's) (1)	Absolute values of differences (1000's) (2)	Absolute % differences (3)	Index of dissimilarity (4)
1	4 region model	-397	1,854	1.91	0.22
2	components of growth	57,253	57,253	112.07	1.59
3	bi-regional aggregation	-6,675	7,699	16.01	0.74
4	single region, net migration	-28,920	54,700	108.74	6.83
5	decomposition A with net migration	-19,590	37,060	77.32	4.76
6	decomposition B with net migration	-13,373	25,665	53.37	3.28
7	single region with components of growth	2,689	13,637	28.55	1.58
8	decomposition B with bi-regional aggregation	-1,308	5,432	10.94	0.62
Meaning of statistic		US under or over prediction	Total deviations	Equal weights for each region	Relative distribution statistic

Source: Computed from the projected population tables in Rogers (1976).
Definitions of comparison statistics:
P(i,m) = population of region i in 2008 projected by model m.
P(*,m) = population of US in 2008 projected by model m.
m = 9 = 9 x 9 multiregional model.

(1) Difference in US total population = P(*,9) - P(*,m).
Negative values indicate overprediction, positive underprediction.
(2) Absolute value of differences
= sum for all regions of absolute values of (P(i,9) - P(i,m)).
(3) Absolute percentage difference
= sum for all regions of absolute values of (P(i,9)-P(i,m))/P(i,9).
(4) Index of dissimilarity
= sum for all regions of absolute values of
((P(i,9)/P(*,9)) - (P(i,m)/P(*,m))).

loss or when natural loss is balanced by net immigration gain.

Empirically, for very many countries, international migration flows are a very important component of population change. It has been estimated (Rees and Willekens, 1981, Table 8) that immigration flows from outside the country make up about 32 per cent of total in-migration flows into both Dutch and British regions, and 46 per cent of such flows into Canadian provinces. Note that in the United Kingdom, whereas the overall balance of external migration was very small in relation to other population changes, for individual regions the external balance is a very significant contributor to population change. Greater London, for example, is a substantial net importer of external migrants.

A commonly voiced objection to the full and explicit incorporation of external migration flows in multiregional population projection models is that data on such flows are unavailable (particularly on emigration) or unreliable. However, in countries in which migration data derive from population registers immigration and emigration are carefully monitored. In countries dependent on periodic censuses for migration data, immigration data are as available and reliable as internal migration data. Emigration can be estimated very often from survey data or as a residual flow if population stocks at the start and end of the base period are known. There is thus no longer any excuse for neglecting the issue of system closure, that is, how external migration flows should be incorporated in the projection model.

There are, however, a variety of different ways in which external migration flows should be handled in regional projection models, which raise the same concerns that were addressed by Rogers (1976) with respect to internal migration. Should gross or net migration terms be used, and should these be handled using flows or rates in the model? For a number of reasons, the answers to these questions are a little different from those regarding internal migration terms, as will be explained later.

2.5 Base period accounts

The fifth source of variation identified above, namely the way in which the data for the projection base period are assembled, concerns the interface between the available demographic data and the projection model. The author has long argued that such an interface should take the form of a set of demographic accounts, appropriately tailored to the nature of the data and population system being studied. The role of migration data is crucial here. Three kinds of migration are employable in multiregional projections:

(i) migrants/transitions;

(ii) migrations/moves;

(iii) last residence migrants/last residence migrations.

For each type there exists a different accounting framework. If you have census based migrant data to hand you must construct transition accounts; if your data are moves based on counts of changes of address reported to a registration system, however, you must construct <u>movement</u> accounts. If your data consist of migration measured in a census or survey by place of last residence, as in many Third World countries, you must build <u>last residence</u> accounts, although the theory and models for this type of accounts have yet to be worked out. The empirical applications reported in this paper use movement accounts.

There are a number of choices to be made in constructing sets of population accounts, depending principally on what kind of population stocks data are available and whether independent estimates of emigration are to hand.

In the next section of the paper, the accounting framework to be used with moves data is outlined, and compared with the accounts hitherto developed for transitions data.

3. ACCOUNTING FRAMEWORKS FOR POPULATION PROJECTION

3.1 Transition accounts

Transition accounts are briefly described as they are precursors to and contrast with the movement accounts used later in the empirical applications.

The theory (Rees and Wilson, 1977) and practice (Rees, 1981) of multiregional demographic accounting has to date been developed using census migration data which are counts of transitions between regions over a fixed time interval, as in Figure 2.2. The accounts variables are positioned and defined in Table 2. Note that, because a type II, projection cohort age-time observation plan (as in Figure 1.2) has been adopted the accounts table is quite simple in structure. In order to fill in accounts tables for all cohorts the researcher enters known items in the table and selects an appropriate set of equations for estimating missing elements, principally the existence-death transitions (see Rees, 1981 for operational details). The models used for estimation may very easily be used for projection or the rates needed for a multiregional cohort survival model can be estimated from the accounts as

$$h_x^{ei,sj} = K_x^{ei,sj} / K_x^{ei,..} \quad (1).$$

3.2 Movement accounts

Although rates generated from multiregional transition accounts match

TABLE 2. A transition accounts table for cohort x.

Initial state in period		Final state in period: Survival at time t+T: Internal regions 1	2	...	N	Outside world 0	Death in period t to t + T: Internal regions 1	2	...	N	Outside world 0	Totals
Internal regions	1	$K_x^{e1,s1}$	$K_x^{e1,s2}$	...	$K_x^{e1,sN}$	$K_x^{e1,s0}$	$K_x^{e1,d1}$	$K_x^{e1,d2}$	...	$K_x^{e1,dN}$	$K_x^{e1,d0}$	$K_x^{e1,\cdot\cdot}$
	2	$K_x^{e2,s1}$	$K_x^{e2,s2}$	...	$K_x^{e2,sN}$	$K_x^{e2,s0}$	$K_x^{e2,d1}$	$K_x^{e2,d2}$	...	$K_x^{e2,dN}$	$K_x^{e2,d0}$	$K_x^{e2,\cdot\cdot}$
	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$			$\vdots$	$\vdots$	$\vdots$
	N	$K_x^{eN,s1}$	$K_x^{eN,s2}$	...	$K_x^{eN,sN}$	$K_x^{eN,s0}$	$K_x^{eN,d1}$	$K_x^{eN,d2}$	...	$K_x^{eN,dN}$	$K_x^{eN,d0}$	$K_x^{eN,\cdot\cdot}$
Outside world	0	$K_x^{e0,s1}$	$K_x^{e0,s2}$	...	$K_x^{e0,sN}$	$\emptyset$	$K_x^{e0,d1}$	$K_x^{e0,d2}$	...	$K_x^{e0,dN}$	$\emptyset$	$K_x^{e0,\cdot\cdot}$
Totals	.	$K_x^{e\cdot,s1}$	$K_x^{e\cdot,s2}$	...	$K_x^{e\cdot,sN}$	$K_x^{e\cdot,s0}$	$K_x^{e\cdot,d1}$	$K_x^{e\cdot,d2}$	...	$K_x^{e\cdot,dN}$	$K_x^{e\cdot,d0}$	$K_x^{e\cdot,\cdot\cdot}$

Notes:

1. Infant accounts have the same structure: a b superscript is substituted for e, and the value of x set to -n.
2. The x subscript means "aged x to x + n at time t, aged x + n to x + 2 at time t + T for survivor variables; the subscript means "aged x to x + n at time t, dying before attainment of age x + n at time t + T for non-survivor variables.

Definitions of variables in TABLE 2.

t = time at start of period.

T = length of period.

$t+T$ = time at end of period.

K = count of transitions or persons making a transition.

$K_x^{ei,sj}$ = persons in cohort x in existence in region i who survive in region j at time t+T.

$K_x^{ei,dj}$ = persons in cohort x in existence in region i who die in region j before time t+T.

$K_x^{ei,..}$ = population aged x to x+n of region i at time t.

$K_x^{..,dj}$ = total deaths in cohort x in region j in time interval.

$K_x^{..,sj}$ = population aged x+n to x+2n of region j at time t+T.

$K_x^{e0,..}$ = total existing immigrants in cohort x in time interval.

$K_x^{..,s0}$ = total surviving emigrants to all internal regions in time interval.

$K_x^{..,d0}$ = total non-surviving emigrants from all internal regions in time interval.

Where 0 is substituted for i this means the variable refers to external transitions to or from the outside world, 0.

exactly the rates needed in multiregional projection models, they have two disadvantages from a practical point of view. The first is that, in countries where census migration tables of the transition type are available, they are produced intermittently at lengthy intervals. The second is that in many countries movement data are gathered via registers and no migration question is used in the national census.

The movement accounts table that uses counts of moves between regions over a fixed time interval (as in Figure 2.3) in a type II, projection cohort, age-time observation plan (Figure 1.2) is set out in Table 3. The structure of the table is remarkably simple. Inter-regional mortality transitions which occupied an N region by N region matrix in the Table 2 transition accounts become a column vector of regional deaths; the two column vectors of surviving and non-surviving emigrants become one vector of regional emigrations; the two row vectors of surviving and non-surviving immigrants become one row vector of regional immigrations; the N by N matrix of inter-regional surviving migrants and surviving regional stayers becomes an N by N matrix with interregional migrations in the off diagonal elements and accounting balances in the principal diagonal. The initial and final population vectors in the movement accounts are exactly equivalent to the initial and final population stocks in the transition accounts. The column vector of regional deaths in the movement accounts is also equivalent to the row vector of total regional deaths in the

TABLE 3. A movement accounts table for cohort x in one time interval.

State before move		State after move: Destinations, Internal regions 1	2	...	N	Outside 0	Death	Totals
Origins, Internal regions	1	R^{1}_{x}	M^{12}_{x}	...	M^{1N}_{x}	M^{10}_{x}	D^{1}_{x}	$P^{1.}_{x}$
	2	M^{21}_{x}	R^{2}_{x}	...	M^{2N}_{x}	M^{20}_{x}	D^{2}_{x}	$P^{2.}_{x}$
	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
	N	M^{N1}_{x}	M^{N2}_{x}	...	R^{N}_{x}	M^{N0}_{x}	D^{N}_{x}	$P^{N.}_{x}$
Outside world		M^{01}_{x}	M^{02}_{x}	...	M^{0N}_{x}	0	0	$M^{0.}_{x}$
Totals		$P^{.1}_{x}$	$P^{.2}_{x}$	...	$P^{.N}_{x}$	$M^{.0}_{x}$	$D^{.}_{x}$	$T^{..}_{x}$

Notes:

1. The infant accounts, for x = −n, follow the same structure except that variables B^{1} ... B^{N} replace $P^{1.}_{-n}$ ... $P^{N.}_{-n}$.
2. The x subscript refers to a "projection cohort" occupying the age-time space x,t; x+n,t; x+n,t+T; x+2n,t+T in Figure 1 (type II observation plan).

Definitions of variables in TABLE 3.

POPULATION

$P_x^{i.}$ = population in region i aged x to x+n at start of time interval.

$P_x^{.j}$ = population in region j aged x+n to x+2n at end of time interval.

DEATHS

D_x^{i} = deaths in region i to persons in cohort x during time interval.

$D_x^{.}$ = total deaths in all internal regions.

EMIGRATIONS

M_x^{iO} = emigrations from region i by persons in cohort x to the outside world, O, during the time interval.

$M_x^{.O}$ = total emigrations from all internal regions.

INTERNAL MIGRATIONS

M_x^{ij} = migrations from region i to region j by persons in cohort x during the time interval.

RESIDUAL TERM

R_x^{i} = residual accounts balancing term for region i.

IMMIGRATIONS

M_x^{Oj} = immigrations from the outside world, O, to region j by persons in cohort x during the time interval

$M_x^{O.}$ = total immigrations to all internal regions.

$T_x^{..}$ = total of all inflows/inputs to the internal system
= total of all outflows/outputs from the internal system.

transition accounts.

An extract from a movement accounts table for a 20 zone UK system is given in Table 4. Consider the row of the table for the Central Clydeside Conurbation (the City of Glasgow and adjacent districts). The region starts with a mid-year 1976 population aged 20-24 of 124,800. This population is subjected to the attrition of some 487 deaths, 8,278 emigrations, 391 outmigrations to Wales, 885 to the North West Remainder and so on including 16,200 outmigrations to Scotland Remainder and 304 to Northern Ireland over the period 1976-81. The balance element in Central Clydeside is 83,820 - note that this is not the number of persons who do not move over the period nor the number of stayers. It is the number of interregional stayers less the number of interregional out-migrations surplus to the number of transitions for the region in question! To the balancing element are added in-migrations of 261 from Northern Ireland, 9,115 from Scotland Remainder ... 490 from the North West Remainder and 278 from Wales together with immigrations from overseas of 2,552, leaving some 103,237 persons aged 25-29 living at mid-1981 in Central Clydeside.

The figures can be reassembled in simpler form, region by region, for components of change in gross and net terms. Table 5 shows the components of change for Central Clydeside. Internal out-migrations totalled 32,215 and the counterstream was only 16,865. The net external migration balance was -5,726 and the net internal balance -15,350; total net out-migration

TABLE 4. An extract from a movement accounts table for a 20 zone UK system, 1976-81, for cohort 20-24 to 25-29

State after move / State before move		Destination in UK NI	CC	SR	...	NWR	W	Emigrations	Deaths	Totals
Northern Ireland	NI	103466	261	631	...	493	288	2363	591	115400
Central Clydeside	CC	304	83820	16200	...	885	391	8278	487	124800
Scotland Remainder	SR	714	9115	196429	...	1636	926	10643	994	250300
.		.	.	.	.	.	.	.	.	.
.		.	.	.	.	.	.	.	.	.
.		.	.	.	.	.	.	.	.	.
North West Remainder	NWR	377	490	1282	...	101801	2782	5482	552	149000
Wales	W	215	278	945	...	2464	142560	6782	615	186500
Immigrations		1289	2552	6795	...	5678	3848	0	0	167811
Totals		112139	103237	247354	...	153872	182940	189011	13037	4039011

TABLE 5. <u>Components of change for a region, Central Clydeside, 1976-81</u>

Gross component	Gross flows	Net flows	Net component
Initial population	124,800	124,800	Initial population
Deaths	− 487	− 487	Deaths
Emigrations	− 8,278	− 5,726	Net external immigrations
Total internal out-migrations	− 32,215	− 15,350	Net internal in-migrations
Residuals	= 83,820	(− 21,076	Net migrations)
Total internal in-migrations	+ 16,865		
Immigrations	+ 2,552		
Final population	= 103,237	= 103,237	Final population

was 21,076. Central Clydeside was not a popular region for those in their twenties in the later 1970's and early 1980's.

From the transition accounts table, matrices of inter-regional survivorships rates can be computed by dividing elements in the existence-survival half of the table by their population at risk, the initial populations of the origin region. These rates can be entered directly into a multiregional cohort survival model projection.

For movement accounts, the population at risk is the average population in the period, which can most conveniently be estimated as half of the initial and final populations. The rates can thus be computed as, using lower case letters for rates:

$$m_x^{ij} = M_x^{ij} / 0.5(P_x^{i.} + P_x^{.i})$$

$$m_x^{i0} = M_x^{i0} / 0.5(P_x^{i.} + P_x^{.i}) \qquad (2)$$

These are transmission rates because the origin region population at risk are used. Death rates are, as conventionally:

$$d_x^{i} = D_x^{i} / 0.5(P_x^{i.} + P_x^{.i}) \qquad (3)$$

but for immigrations we must use the population at risk of the destination regions since the origin region (the outside world) does not appear in the accounts. Thus the immigration rates are admission rates:

$$m^{Oi}_x = M^{Oi}_x / 0.5(P^{i.}_x + P^{.i}_x) \qquad (4)$$

Table 6 shows the rates equivalent to the flows previously displayed in Table 4.

3.3 <u>A projection model based on movement accounts</u>

How should these movement rates be employed in a multiregional population projection model? It turns out that all that is needed is to rearrange the rate equations into the equivalent flow equations:

$$M^{ij}_x = m^{ij}_x \, 0.5(P^{i.}_x + P^{.i}_x)$$

$$M^{iO}_x = m^{iO}_x \, 0.5(P^{i.}_x + P^{.i}_x)$$

$$D^{i}_x = d^{i}_x \, 0.5(P^{i.}_x + P^{.i}_x)$$

$$M^{Oi}_x = m^{Oi}_x \, 0.5(P^{i.}_x + P^{.i}_x) \qquad (5)$$

These are the items we need to enter in the accounts table to estimate the residual balancing terms

$$R^{i}_x = P^{i.}_x - D^{i}_x - M^{iO}_x - \sum_{j=i} M^{ij}_x \qquad (6)$$

and the final populations

TABLE 6. <u>An extract from a movement rates table for a 20 zone UK system, 1976-81, for cohort 20-24 to 25-29</u>

State before move / State after move		Destination in UK: NI	CC	SR	...	NWR	W	Emigrations	Deaths	Populations at risk
Northern Ireland	NI	–	.00229	.00555	...	.00433	.00253	.02077	.00519	113,769.5
Central Clydeside	CC	.00267	–	.14208	...	.00776	.00343	.07260	.00427	114,018.5
Scotland Remainder	SR	.00287	.03663	–	...	.00657	.00372	.04277	.00399	248,827
.		.	.	.	.	.	.	.	.	.
.		.	.	.	.	.	.	.	.	.
.		.	.	.	.	.	.	.	.	.
North West Remainder	NWR	.00249	.00324	.00847	...	–	.01837	.03620	.00345	151,436
Wales	W	.00116	.00150	.00512	...	.01334	–	.03672	.00333	184,720
Immigrations		.01133	.02238	.02731	...	.03749	.02083	–	–	.04354
Totals		-.02866	-.18912	-.01184	...	.03217	-.01927	.04904	.00338	3,854,081.5

Notes:
1. Rates in the NI to W rows are transmission rates, flows/origin population at risk.
2. Rates in the immigrations row are admission rates, flows/destination population at risk.
3. Rates in the totals row are population change/population at risk, except for the emigrations and deaths columns which contain UK rates. The toal UK change rate = -.00888.

$$P_x^{.i} = R_x^i - \sum_{j=i} M_x^{ji} + M_x^{Oi} \qquad (7)$$

But, you object, you can't use the equation (5) set because you don't know the final populations until you get to equation (7). Well simply enter them as zero the first time equation set (5) is computed, work out an initial estimate of the final populations and then return with this estimate to equation (5). Repeat this sequence of calculations until your population at risk estimates don't change very much any more, and you will then have projected the population forward one time interval. These iterative calculations can be speeded up by using the initial populations as populations at risk on the first iteration through the equations.

The infant cohort (persons born during the time interval) should be computed after all the "existing" cohorts have been projected. Births will simply be projected as

$$B_x^i = b_x^i \; 0.5(P_x^{i.} + P_x^{.i}) \qquad (8)$$

where B_x^i are births in region i to mothers in projection cohort x for all fertile cohorts, say x = 10-14 to 15-19 through 40-44 to 45-49, and then summed

$$B_{\cdot}^i = \sum_x B_x^i \qquad (9)$$

These are then equivalent to the row total in the infant cohort accounts, that is

$$P^{i}_{-n} = B^{i} \tag{10}$$

and the projection model above is used again.

This structure of equations is very flexible in that we can easily alter our estimates for any components. For example, we could assume immigrations to be fixed by legal quota

$$M^{Oi}_{x} = M^{Oi}_{x} \text{ (fixed)} \tag{11}$$

3.4 <u>A multiregional cohort survival model</u>

Is iteration through the projection model equations strictly necessary? In certain circumstances, when rates are assumed constant over the whole of the projection period, it may be convenient to substitute an analytical solution to the projection equations for the iterative. All we have to do is to substitute equations (7) and (6) back into equation set (5) to yield rate equations of the kind:

$$M^{ij}_{x} = m^{ij}_{x} \,/\, 0.5(P^{i.}_{x} + ((P^{i.}_{x} - D^{i}_{x} - M^{iO}_{x} - \Sigma_{j=i} M^{ij}_{x} + \Sigma_{j=i} M^{ji}_{x} + M^{Oi}_{x})) \tag{12}$$

The problems with this equation are that the migration terms appear on the right hand side of the equation as well as the left, and in two guises, in the out-migration and in-migration sums, and that the equation for one inter-regional migration flow is dependent on the others. We need to rearrange terms such that we have equations of the form

$$\text{final population} = \text{rates} \times \text{initial population} \tag{13}$$

It is easiest to do this using a matrix formulation. The task is to define $\underline{S}_x$ such that

$$\underline{P}_x (t+T) = \underline{S}_x \underline{P}_x (t) \tag{14}$$

where $\underline{P}_x$ (t+T) is a vector of final populations of regions in cohort x at time t+T, $\underline{P}_x$ (t) is a vector of initial populations in cohort x, $\underline{S}_x$ is the rates matrix that transforms the initial population vector into the final.

The matrix of S rates can be defined in a way similar to that of the matrix of P rates, survival probabilities, in life table analysis (Rogers and Ledent, 1976; Willekens and Rogers, 1978).

Let $\underline{M}_x$ be a matrix of migration and mortality rates arranged in the following way, keeping to the definition of x as a "projection cohort":

$$\underline{M}_x = \begin{bmatrix} (d_x^1 + \sum_{j=i} m_x^{1j} + m_x^{10} - m_x^{01}) & -m_x^{21} & \cdots & -m_x^{N1} \\ -m_x^{12} & (d_x^2 + \sum_{j=i} m_x^{2j} + m_x^{20} - m_x^{02}) & \cdots & -m_x^{N2} \\ \vdots & \vdots & & \vdots \\ -m_x^{1N} & -m_x^{2N} & \cdots & (d_x^N + \sum_{j=i} m_x^{Nj} + m_x^{N0} - m_x^{0N}) \end{bmatrix} \tag{15}$$

Then,

$$\underline{S}_x = \{\underline{I}_x + 0.5\underline{M}_x\}^{-1} \{\underline{I}_x - 0.5\underline{M}_x\} \qquad (16)$$

Equation (15) extends the previous formulation by adding emigration rates to the diagonal term in the $\underline{M}$ matrix and subtracting immigration rates.

If we substitute equation (16) into the RH side of equation (14) we now have a one equation projection model based on movement accounts:

$$\underline{P}_x(t+T) = \{\underline{I} + 0.5\underline{M}_x\}^{-1} \{\underline{I} - 0.5\underline{M}_x\} \underline{P}_x(t) \qquad (17)$$

This is convenient for the analytical exploration of the consequences of the current pattern of multiregional population change (stable growth analysis), but it involves committment to only one of the methods for handling external migration. Since this is one of the key issues addressed in this paper the more flexible projection model is used.

It should be stressed that the matrix of survivorship rates defined in equation (16) is not a matrix of transition probabilities. They do not yield the probability that a person living in one region at the start of a period will be in another at the end. They are merely those rates, based on observed mortality and mobility, that transform the initial into the final population distribution across regions.

4. ACCOUNTS BASED MODELS: THE WAYS IN WHICH THE DATA ARE ASSEMBLED

There are a variety of ways of putting together the components of multiregional change to form population accounts, the basis of projections. Figure 3 sets out this variety. The top diagram in Figure 3 names the components of movement accounts that have already appeared in algebraic and numerical form. To construct accounts tables it is essential to know the regional distribution of mortality, fertility and immigration and the interregional distribution of internal migration. Thus, in all four alternative ways of assembling accounts (Figures 3.1 through 3.4) these components are numbered and shaded. It is necessary, in addition, to be able to estimate at least two of the following components: initial regional populations, final regional populations and emigration flows from the regions. Combinations of these components lead to four types of model for constructing population accounts.

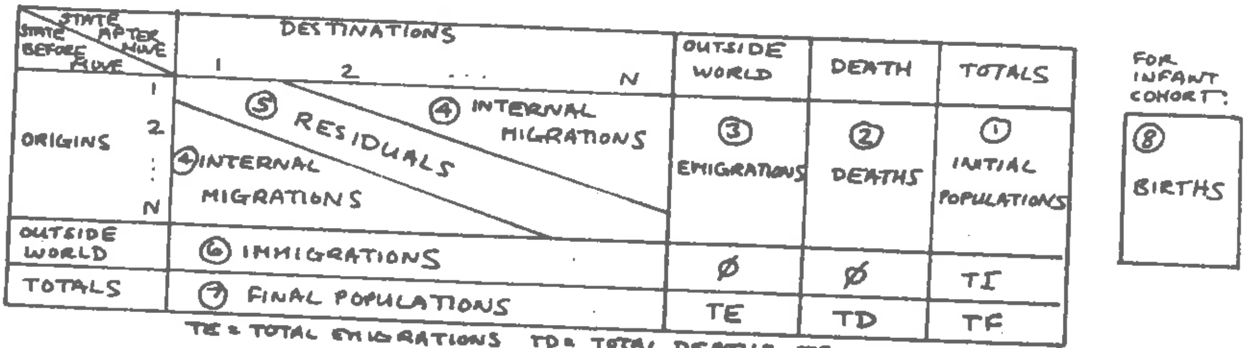
4.1 The forecast model

Here initial populations and emigrations are used as inputs and final populations are generated from the accounts as the totals for the columns. Figure 3.2 shows the forecast model graphically.

4.2 The backcast model

For periods ending with a decennial census it may be preferable to use the final populations, reliably estimated from a recent census rather than the initial populations, along with emigration flows. This is the backcast

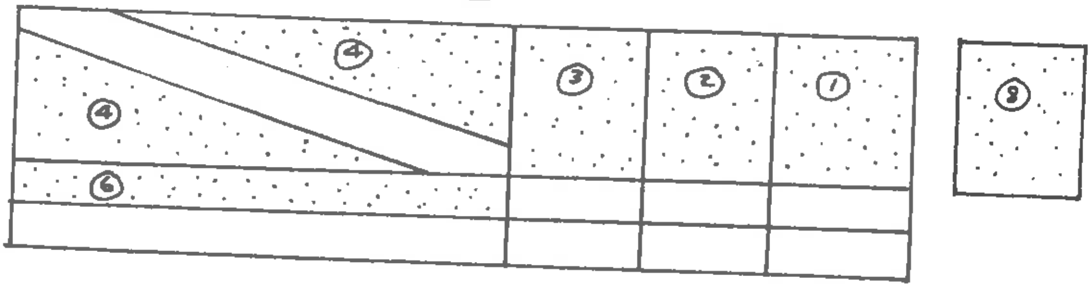
THE COMPONENTS



TE = TOTAL EMIGRATIONS TD = TOTAL DEATHS TF = TOTAL FLOWS TI = TOTAL IMMIGRA

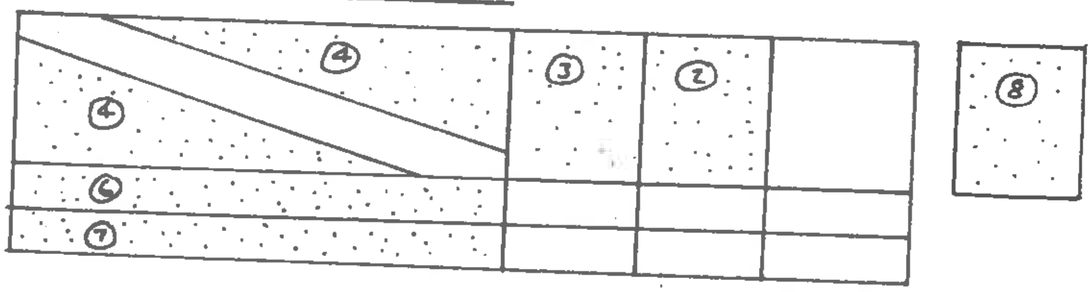
3.1 THE FORECAST MODEL : DATA INPUTS

CHOICE 1



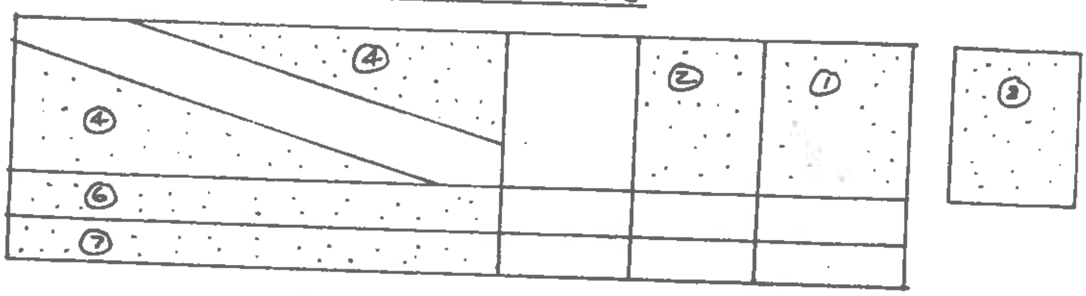
3.2 THE BACKCAST MODEL : DATA INPUTS

CHOICE 2



3.3 THE NO-EMIGRATION MODEL : DATA INPUTS

CHOICE 3



3.4 THE BOTH POPULATIONS MODEL : DATA INPUTS

CHOICE 4

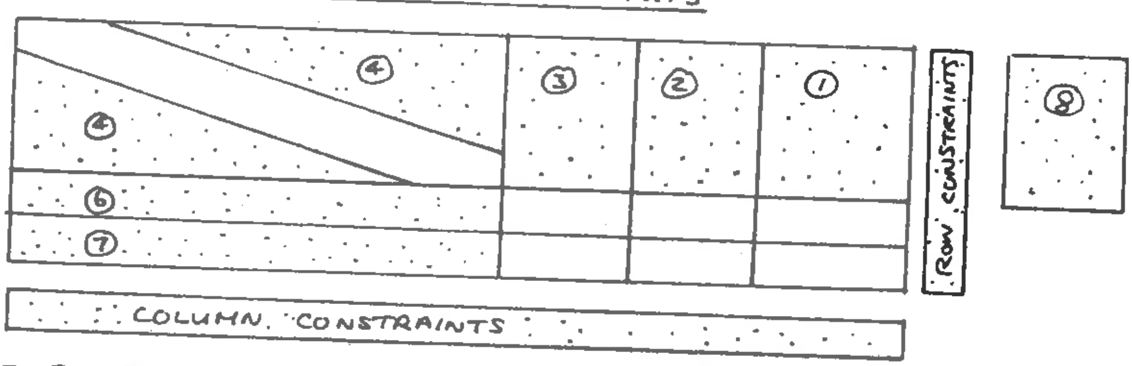


FIGURE 3. THE DIFFERENT WAYS OF ASSEMBLING MOVEMENT ACCOUN

model exhibited in Figure 3.3. From this model the initial populations are estimated by adding across the rows of the accounts table.

4.3 The no-emigration model

In certain situations emigration may not be available or may be very unreliable. Then emigration flows can be estimated as residuals if both initial and final populations are input, as in Figure 3.4. However, there is the problem with this model that all data errors pile up in the emigration terms , and in some instances negative estimates may result. The no-emigration model is most appropriately used with transition data that links two successive censuses, as is the situation in Canada and in France. When movement data can be used reasonable estimates of external moves can be made, and this model will not be used. The model is not employed in the projection experiments described later in the paper.

4.4 The both populations model and constraints

All eight components of the movement accounts can be introduced into the tables. However, in this case there will be inconsistency either between the final populations computed from the columns of the table and those input as data, when the estimates are based on a forecast model, or between the initial populations computed from the rows of the table and those input as data, when the estimates are based on a backcast model.

The solution is to adjust the estimates of a forecast or backcast model to independent sets of row and column constraints containing the initial and final population vectors, emigration total, deaths total and immigration total, modified to ensure equality of the sum of row and column totals. If A(i,j) are defined to be the estimates of accounts elements in row i and column j in each cohort table, and R(i) and C(j) are the independent row and column constraints, then constrained re-estimates of the accounts elements are obtained by iteration through the following sets of equations:

$$A(i,j) = A(i,j) \{ R(i)/\sum_j A(i,j) \} \qquad (18)$$

$$A(i,j) = A(i,j) \{ C(j)/\sum_i A(i,j) \} \qquad (19)$$

until all elements differ by less than half a person from one iteration, p-1, to the next, p:

$$\left| A^{p}(i,j) - A^{p-1}(i,j) \right| < 0.5 \qquad (20)$$

Note that adjustment to constraints is only possible if

$$\sum_i R^{i} = \sum_j C^{j} \qquad (21)$$

and some of the constraints must be adjusted if this is not true.

The empirical experiments described later are designed to measure the degree to which the base period accounts and projected populations for a 20 zone UK system differ according to the way in which the data are assembled.

5. THE WAYS IN WHICH THE POPULATION SYSTEM IS CLOSED

The second set of choices to be explored empirically relate to the ways in which a system of regions is closed. Here we assume the system consists of a number of regions of interest together with the rest of the country.

5.1 External migration ignored

The first choice is to ignore external migration. This assumes the country is a closed system. But the system will still be influenced by external flows if surviving stayer (transition accounts) or residual (movement accounts) terms are computed as one minus the sum of regional internal out-migration rates and minus the regional death rates. In these cases the surviving stayer flows or residual elements will be overestimated by the sizes of the emigration flows; this error will be partly compensated for by the absence of any immigration terms.

Figure 4 sets out the closure choices in diagramatic fashion. Figure 4.1 shows how this first choice is implemented in the movement accounts model: all emigrations and immigrations are entered as zero flows.

5.2 Net migration flows or rates

The second choice is to incorporate net external migration as net flows or rates. One of these alternatives has been the usual choice applied in national population projections: those of the United States or United Kingdom use vectors of net immigration flows. Academic forecasters

FIGURE 4. THE DIFFERENT WAYS OF CLOSING THE SYSTEM USING MOVEMENT ACCOUNTS

THE COMPONENTS OF THE MOVEMENT ACCOUNTS

STATE BEFORE MOVE / STATE AFTER MOVE		DESTINATIONS 1 2 ... N	OUTSIDE WORLD	DEATH	TOTALS
ORIGINS	1 2 : N		EMIGRATIONS		
OUTSIDE WORLD		IMMIGRATIONS			
TOTALS					

4.1 EXTERNAL MIGRATION IGNORED

CHOICE 1	ZEROES
ZEROES	

4.2 NET MIGRATION FLOWS OR RATES

CHOICE 2.1	ZEROES
NET IMMIGRATION FLOWS	

CHOICE 2.2	ZEROES
NET IMMIGRATION RATES (A)	

4.3 GROSS MIGRATION FLOWS OR RATES

CHOICE 3.1	EMIGRATION FLOWS
IMMIGRATION FLOWS	

CHOICE 3.2	EMIGRATION RATES (T)
IMMIGRATION RATES (A)	

4.4 MIXED GROSS MIGRATION FLOWS AND RATES

CHOICE 4.1	EMIGRATION RATES (T)
IMMIGRATION FLOWS	

CHOICE 4.2	EMIGRATION FLOWS
IMMIGRATION RATES (A)	

4.5 A WORLD SYSTEM

CHOICE 5	EMIGRATION RATES (T)
REST OF WORLD OUT-MIGRATION RATES (T)	

(A) = ADMISSION RATES

(T) = TRANSMISSION RATES

prefer to use net immigration (admission) rates. The multiregional cohort survival model based on movement rates incorporated net external migration rates (equation 15). Figure 4.2 shows that these are introduced into the movement accounts model as net immigration flows or rates, which may be positive or negative, and emigrations are input as zeroes.

This second choice improves on the first but has similar disadvantages to those of conventional net migration models as identified by Rogers (1976). A net migration is not a move and so cannot be traced in the projection.

5.3 <u>Gross migration flows or rates</u>

The third choice is to introduce external migrants as gross flows or rates (Figure 4.3). This does not produce very different results from the corresponding net flow or rate model if the corresponding gross and net flows are used, but it is easier in the gross version to change the separate flows or rate components as immigration and emigration are subject to rather different influences.

Although in modelling multiregional population systems internal to a country it is conventional to use migration rates rather than flows so that the projected migration streams can respond to changes in origin region population, there may often be a case for using flows. The reason is that external migration, particularly immigration, is subject to much more government control than is internal migration.

5.4 Mixed gross migrations flows and rates

In fact, it may be most appropriate to project emigration using rates because the controls are distant and diverse (in other countries) and to project immigration as a flow subject to national control, as in Figure 4.3. The obverse choice - immigration rates and emigration flows - is probably less appropriate.

5.5 A world system

The final choice would be to include the rest of the world simply as another region in the multiregional system. It is probably unrealistic to project external flows on this basis since immigration flows into a country will increase as rapidly as the world population (currently about 1.5 per cent per annum), but it is interesting to have this projection as a benchmark to measure the effect of immigration control on regional and national populations.

6. RESEARCH DESIGN

To assess the effect of the two choice dimensions described in sections 4 and 5 a research design was set up and some eight different projections of the UK population were carried out. Since Figure 3 lists 4 choices and Figure 4 has 8 choices a full matrix of projections of every choice pair would involve 32 projections. These were whittled down to eight in the following way.

The no-emigration model was not investigated since for the UK demographic information system used immigration and emigration were equally available. The different accounts models were all run in combination with closure choice 3.2, involving emigration rates and immigration flows.

From the system closure alternatives, choice 1, involving no external migration, choice 2.2, involving net immigration rates, choice 3.1, involving gross migration flows, choice 3.2, involving gross migration rates and immigration flows and choice 5, involving a world system were selected, all in combination with the forecast model option. The result is eight projections , shown in Table 7, which cover both choice dimensions. However, the assumption is that there are no interaction effects in the table between the dimensions: this hypothesis should be tested in further analyses.

TABLE 7. The research design

Ways of closing the system	Ways of assembling movement accounts		
	Choice 1: Forecast model	Choice 2: Backcast model	Choice 4: Constrained model
Choice 1: External migration ignored	Projection model 4		
Choice 2.2: Net migration rates	Projection model 5		
Choice 3.1: Gross migration flows	Projection model 6		
Choice 3.2: Gross migration rates	Projection model 1	Projection model 2	Projection model 3
Choice 4.1: Emigration rates Immigration flows	Projection model 7		
Choice 5: A world system	Projection model 8		

7. A DEMOGRAPHIC INFORMATION SYSTEM FOR 20 UK ZONES

To gauge the effect of data assembly and system closure choices, it was necessary to put together the components of a set of movement accounts for a real world system. This turned out to be a very extensive task, of which only the barest bones are sketched here. A full report will be prepared in another paper.

7.1 Zones

The zone system consists of 20 units for which demographic data were published with some degree of age disaggregation. Figure 5 shows the boundaries of the zones which completely cover the United Kingdom. Eight zones are metropolitan counties or their administrative or statistical equivalents; seven zones consist of the whole or part of the rest of a standard region (that is, the region minus the metropolitan zone); and five zones are standard regions in which a metropolitan core is not distinguished. A previous paper (Rees and Stillwell, 1982) showed that the main contrasts in population change in this 20 zone system were between metropolitan and non-metropolitan zones rather than between northern and southern. This spatial disaggregation is thus a considerable improvement over previously used standard region systems though it is still not as fine as the functional urban region system developed by the Centre for Urban and Regional Development Studies at the University of Newcastle. Construction of multiregional demographic accounts and projections for this latter

FIGURE 5. The study set of zones

system was felt to be a little too ambitious an undertaking for present purposes.

7.2 Time periods

The time period chosen was the interval mid-year 1976 to mid-year 1981. A five year period was needed to match the anticipated level of reliable age disaggregation (five year age intervals). But most importantly movement data for the UK have only recently become available from mid-1975. The end-point of the period, mid-year 1981, comes just after the 1981 Census, and the 1981 population estimates thus provide a sound basis on which to project the population.

7.3 Age classifications

The age classifications used in the various component data sets proved to be a veritable jungle to wade through. Table 8 sets out the various age disaggregations in terms of best, worst or standard cases.

The age classification of a component was rarely consistent across all zones and all years. Three different "national" agencies provide demographic data for England and Wales, Scotland and Northern Ireland respectively. The best data were generally for Scotland and the worst for Northern Ireland. The age classifications also varied from year to year: the best breakdowns were for 1977 through 1980. They tended to be worse in 1981 because of the time lag in publication or because of a reduction in the number and detail of tables as a result of cutbacks in government

TABLE 8. <u>The age group (type I) classifications of the raw data at zonal level</u>

Populations Best case	Populations Worst case	Deaths Best case	Deaths Worst case	Emigrations & Immigrations National case	Births Standard case	Internal migrations Standard case
0,1-4	0-4	0,1-4	0,1-4			
0-4	0-4	0-4	0-4			0-4
5-9	5-14	5-9	5-14	0-14		5-9
10-14		10-14			<15	10-14
15-19	15-24	15-19	15-24	15-29	15-19	15-19
20-24		20-24			20-24	20-24
25-29	25-34	25-29	25-34		25-29	25-29
30-34		30-34		30-44	30-34	30-34
35-39	35-44	35-39	35-44		35-39	35-39
40-44		40-44			40-44	40-44
45-49	45-54	45-49	45-54	45-59/64	45-49	45-49
50-54		50-54			50+	50-54
55-59	55-64	55-59	55-64			55-59
60-64		60-64		60+/65+		60-64
65-69	65-74	65-69	65-74			65-69
70-74		70-74				70-74
75-79	75+	75-79	75-84			75+
80-84		80-84				
85-89		85-89	85+			
90+		90+				

statistical services.

The best situation was for births: type I data for all zones and for all years for 9 age groups were available. The worst situation was for emigrations and immigrations where the only age disaggregation was for the UK as a whole and was for only five broad age groups.

The Table 8 age classified data (type I) had to be converted into cohort classified data (type II) for the 19 cohorts listed in Table 9.

7.4 The demographic information system

The process by which this and other estimations were achieved is laid out in Figure 6. The process involved 5 sets of about 80 information files (file sets I to V) and 4 sets of about 20 programs (program sets A to D).

Magnetic tapes (files I-M) with the National Health Service Central Register data on transfers between Family Practitioner Areas were supplied by the Office of Population Censuses and Surveys. These data were checked, aggregated and rearranged in a FORTRAN program (Stillwell, 1983) (program A-M) and files referring to the 20 zone UK system created for 1975-76 to 1981-82 (file set II-M). Another FORTRAN program (program B-M) was then used to solve the three face problem (Willekens, Por and Raquillet, 1982) for this data to produce files containing inter-zonal migrations classified by age from the aggregate inter-zonal matrix and age disaggregated zonal in- and out-migrations (file set III-M). These separate year files were then consolidated and the age-time plan converted into cohort form in a Waterloo BASIC program (program C-M).

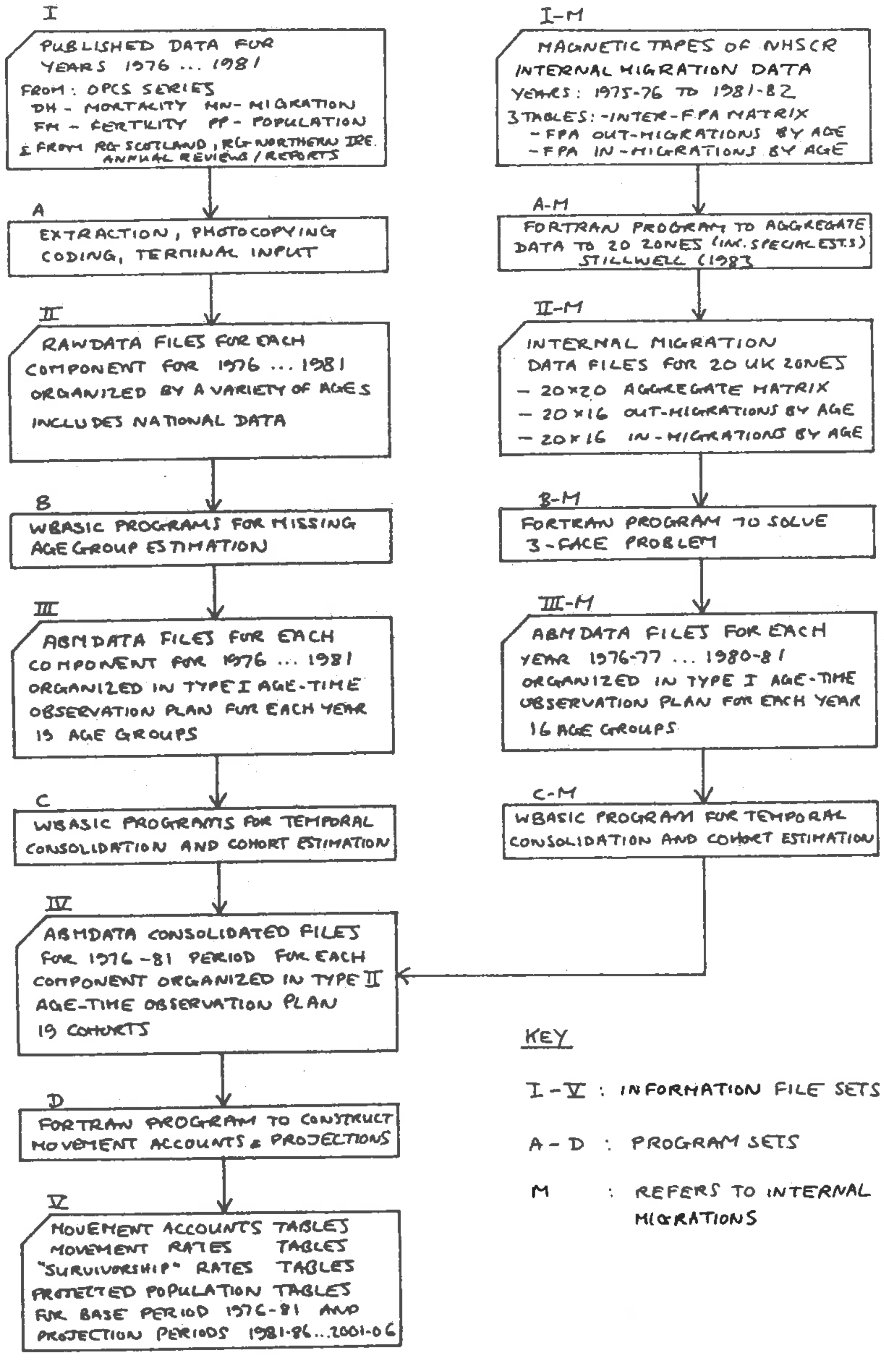
TABLE 9. The cohort (type III) classifications used in the demographic information system

Cohort number	Initial populations	Deaths, emigrations, internal migrations, immigrations and births from	to	Final populations	Cohort order in files
1		birth	0-4	0-4	19
2	0-4	0-4	5-9	5-9	1
3	5-9	5-9	10-14	10-14	2
4	10-14	10-14	15-19	15-19	3
5	15-19	15-19	20-24	20-24	4
6	20-24	20-24	25-29	25-29	5
7	25-29	25-29	30-34	30-34	6
8	30-34	30-34	35-39	35-39	7
9	35-39	35-39	40-44	40-44	8
10	40-44	40-44	45-49	45-49	9
11	45-49	45-49	50-54	50-54	10
12	50-54	50-54	55-59	55-59	11
13	55-59	55-59	60-64	60-64	12
14	60-64	60-64	65-69	65-69	13
15	65-69	65-69	70-74	70-74	14
16	70-74	70-74	75-79	75-79	15
17	75-79	75-79	80-84	80-84	16
18	80-84	80-84	85-89	85-89	17
19	85+	85+	90+	90+	18

Notes:

1. Births data assembled for cohorts 4 through 10, which refer to mothers' cohort at time of maternity.
2. All cohorts are "projection" cohorts, type II in Figure 1.

FIGURE 6. STEPS IN THE ASSEMBLY OF DEMOGRAPHIC DATA FOR THE UK MOVEMENT ACCOUNTS



The other components were processed in the same sequence of operations, although the details differed between components and years. Published data were extracted from OPCS publications (file set I), coded up and input via a VDU terminal to computer files (operation A) to form "rawdata" files (file set II). These "rawdata" files were input to a series of Waterloo BASIC programs (program set B) that estimated the full age disaggregation for age-time plan type I which were written to "abmdata" files (file set III).

These "abmdata" files for six separate calendar years were consolidated into data for the mid-year 1976 to mid-year 1981 period by a series of Waterloo BASIC programs (program set C) and at the same time the data were converted to projection cohort form to produce one data file for each component (file set IV). These files were input to a FORTRAN program (program set D) to construct movement accounts tables , various tables of rates and tables of projected populations for the base period 1976-81 and for five periods of projection 1981-86, 1986-91, 1991-96, 1996-2001 and 2001-06.

8. THE PROJECTION EXPERIMENTS

8.1 General description

Projections were carried out using some eight models (as specified in Table 7). Accounts were constructed for the base period 1976-81, and from this set of accounts the appropriate rates and flows for the projection model were extracted. These were then applied in all models to the 1981 zonal populations, so that each projection began with the same, accurate base population. The projection then rolled forward using the appropriate rates and flows in a constant fashion, applying the rates and/or flows to the projected populations at the end of each projection period. The projections were carried forward over 5 five year periods, a sufficient time span for the present experiments.

Each model provides not only a projection of the populations of each zone by age but also a full set of accounts and components of growth. These are too voluminous to reproduce or discuss here, but will be the subject of future analysis. The projections thus contain forecasts of the volume of internal and external migration, births and deaths as well as of population. Here we concentrate attention on the final projected populations at the end of the projection time span in 2006.

The results for the eight models are collected together in four tables. Table 10 gives the 2006 populations, in millions, for each of the zones as projected by the eight models. Table 11 converts these zonal

TABLE 10. The projected populations of the UK zones in 2006, in millions.

Zone	1981 popn.	Projection model 1 f er,ir	2 b er,ir	3 c er,ir	4 f emi	5 f nmr	6 f ef,if	7 f er,if	8 f world
ni	1.547	1.712	1.720	1.605	1.764	1.712	1.716	1.709	1.723
cc	1.725	1.307	1.244	1.411	1.426	1.307	1.292	1.315	1.341
sr	3.425	3.433	3.551	3.400	3.541	3.433	3.435	3.430	3.503
tw	1.161	0.976	0.967	1.029	1.020	0.976	0.974	0.980	0.997
nr	1.953	1.831	1.865	1.923	1.895	1.831	1.831	1.834	1.867
sy	1.315	1.248	1.268	1.258	1.260	1.248	1.249	1.250	1.274
wy	2.065	1.923	1.942	2.012	1.951	1.923	1.925	1.926	1.964
yhr	1.526	1.548	1.591	1.577	1.555	1.548	1.550	1.547	1.585
em	3.840	3.926	4.054	4.158	3.998	3.926	3.932	3.923	3.995
ea	1.895	2.115	2.227	2.173	2.187	2.115	2.123	2.109	2.159
ose	4.734	5.633	5.986	5.245	5.655	5.633	5.643	5.591	5.767
oma	5.442	5.422	5.594	5.738	5.564	5.422	5.429	5.428	5.613
gl	6.851	6.055	5.942	6.161	5.971	6.055	6.075	6.120	6.510
sw	4.363	4.669	4.815	4.682	4.711	4.669	4.675	4.657	4.771
wmc	2.674	2.293	2.262	2.309	2.300	2.293	2.294	2.304	2.354
wmr	2.507	2.738	2.891	2.692	2.672	2.738	2.737	2.731	2.807
gm	2.624	2.341	2.335	2.349	2.379	2.341	2.341	2.347	2.394
m	1.525	1.269	1.242	1.274	1.298	1.269	1.268	1.276	1.302
nwr	2.315	2.323	2.401	2.300	2.327	2.323	2.325	2.321	2.382
w	2.807	2.751	2.809	2.868	2.818	2.751	2.753	2.751	2.800
UK	56.289	55.515	56.705	56.165	56.294	55.515	55.569	55.548	57.106

Notes:

1. Ways of assembling accounts:
   f = forecast model b = backcast model c = constrained model
2. Ways of closing the system:
   er,ir = emigration rates, immigration rates
   emi = external migration ignored
   nmr = net migration rates
   ef,if = emigration flows, immigration flows
   er,if = emigration rates, immigration flows
   world = world as one internal system
3. Zones:

ni = Northern Ireland cc = Central Clydeside sr = Scotland Remainder
tw = Tyne and Wear nr = Northern Remainder sy = South Yorkshire
wy = West Yorkshire yhr = Yorks.&Humber.Rem. em = East Midlands
ea = East Anglia ose = Outer South East oma = Outer Metro. Area
gl = Greater London sw = South West wmc = West Midlands County
wmr = West Midlands Rem. gm = Greater Manchester m = Merseyside
nwr = North West Rem. w = Wales
See Figure 5 for boundaries. UK = United Kingdom

TABLE 11. The projected population distribution by zone for the UK, 2006. Percentages of the UK population.

Zone	1981 %	Projection model							
		1	2	3	4	5	6	7	8
		f	b	c	f	f	f	f	f
		er,ir	er,ir	er,ir	emi	nmr	ef,if	er,if	world
ni	2.75	3.08	3.03	2.86	3.13	3.08	3.09	3.08	3.02
cc	3.06	2.35	2.19	2.51	2.53	2.35	2.33	2.37	2.35
sr	6.08	6.18	6.26	6.05	6.29	6.18	6.18	6.18	6.13
tw	2.06	1.76	1.71	1.83	1.81	1.76	1.75	1.76	1.75
nr	3.47	3.30	3.29	3.42	3.37	3.30	3.30	3.30	3.27
sy	2.34	2.25	2.24	2.24	2.24	2.25	2.25	2.25	2.23
wy	3.67	3.46	3.42	3.58	3.47	3.46	3.46	3.47	3.44
yhr	2.71	2.79	2.81	2.81	2.76	2.79	2.79	2.78	2.78
em	6.82	7.07	7.15	7.40	7.10	7.07	7.08	7.06	7.00
ea	3.37	3.81	3.93	3.87	3.89	3.81	3.82	3.80	3.78
ose	8.41	10.15	10.56	9.34	10.05	10.15	10.16	10.06	10.10
oma	9.67	9.77	9.86	10.22	9.88	9.77	9.77	9.77	9.83
gl	12.17	10.91	10.48	10.97	10.61	10.91	10.93	11.02	11.40
sw	7.75	8.41	8.49	8.34	8.37	8.41	8.41	8.38	8.35
wmc	4.75	4.13	3.99	4.11	4.09	4.13	4.13	4.15	4.12
wmr	4.45	4.93	5.10	4.79	4.75	4.93	4.93	4.91	4.91
gm	4.66	4.22	4.12	4.18	4.23	4.22	4.21	4.23	4.19
m	2.71	2.29	2.19	2.27	2.31	2.29	2.28	2.30	2.28
nwr	4.10	4.19	4.23	4.09	4.13	4.19	4.18	4.18	4.17
w	4.99	4.95	4.95	5.11	5.01	4.96	4.95	4.95	4.90
UK	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Notes: see Table 10.

populations into percentages of the total UK population. Table 12 provides, for the UK as a whole, the percentage age distribution over 19 age groups projected in 2006 by each model. Finally, Table 13 reports the results of comparing seven of the models with an eighth - model 7 - using the same statistics as were used earlier to evaluate Rogers' results for the USA (Table 1). Model 7, a forecast model using emigration rates and immigration flows, was chosen as the norm for comparison as this was the model that, on the basis of previous arguments would probably have been used for producing population forecasts. In this projection the data for the base period 1976-81 are accepted at face value, and the closure option best reflecting the influences of immigration control on external flows is used.

The results of the model projections are now considered and evaluated.

8.2 What happens to the UK zonal populations?

If we accept, for the moment, that model 7 provides the best projection, what does it tell us will happen to the populations of our 20 UK zones. It firstly suggests that the overall UK population will decline by nearly three quarters of a million from its 1981 value of 56.29 millions, although there will be increases to 1991 and decreases thereafter. The "de-metropolitanization" of the UK continues with all eight metropolitan zones losing population over 1981-2006, particularly Central Clydeside (-24%), Merseyside (-16%) and the West Midlands

TABLE 12. <u>The projected age distribution of the UK population, 2006. Percentages of the UK population.</u>

Age	1981 %	Projection model							
		1 f er,ir	2 b er,ir	3 c er,ir	4 f emi	5 f nmr	6 f ef,if	7 f er,if	8 f world
0-4	6.13	5.86	6.04	5.58	5.80	5.86	5.86	5.93	5.99
5-9	6.54	6.24	6.39	5.92	6.18	6.24	6.24	6.27	6.31
10-14	7.90	6.74	6.90	6.40	6.67	6.74	6.73	6.72	6.73
15-19	8.31	6.86	7.03	6.57	6.79	6.86	6.86	6.87	6.84
20-24	7.45	6.59	6.76	6.28	6.48	6.59	6.58	6.65	6.66
25-29	6.76	6.12	5.99	5.92	6.05	6.12	6.12	6.26	6.32
30-34	7.33	6.45	6.30	6.31	6.46	6.45	6.43	6.58	6.76
35-39	6.40	7.68	7.50	7.65	7.76	7.68	7.69	7.61	7.77
40-44	5.70	7.98	7.81	8.01	8.11	7.98	8.01	7.82	7.93
45-49	5.52	7.03	6.91	7.21	7.19	7.03	7.05	6.95	6.94
50-54	5.71	6.26	6.17	6.47	6.38	6.26	6.26	6.24	6.17
55-59	5.89	6.59	6.50	6.90	6.64	6.59	6.60	6.52	6.41
60-64	5.29	5.42	5.39	5.69	5.42	5.42	5.42	5.40	5.30
65-69	4.99	4.38	4.33	4.58	4.35	4.38	4.37	4.38	4.29
70-74	4.28	3.62	3.52	3.78	3.59	3.62	3.61	3.62	3.54
75-79	3.03	2.89	2.77	3.07	2.87	2.89	2.89	2.89	2.83
80-84	1.72	2.01	1.97	2.11	1.99	2.01	2.01	2.01	1.96
85-89	0.75	0.96	1.08	1.02	0.95	0.96	0.96	0.96	0.94
90+	0.30	0.31	0.66	0.51	0.31	0.31	0.31	0.31	0.30
UK	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Notes: see Table 10.

TABLE 13. Comparisons of the projection models with model 7

Projection number	Model description	Difference in UK total population (1000's) (1)	Absolute value of differences (1000's) (2)	Absolute % differences (3)	Index of dissimilarity (4)
1	f; er,ir	-33	194	6.01	0.18
2	b; er,ir	1,157	1,713	54.20	1.27
3	c; er,ir	617	1,698	60.29	1.53
4	f; emi	746	1,177	46.26	0.74
5	f; nmr	-33	194	6.01	0.19
6	f; ef,if	21	214	7.39	0.21
8	f; world	1,558	1,559	47.80	0.48

Notes:

1. The statistics are computed from the projected figures in Table 10 and 11.
2. The statistics are defined in Table 1.
3. Model descriptions: see Table 10.

Metropolitan County (-14%). Not all the other non-metropolitan zones gain population: Northern Remainder, the Outer Metropolitan Area and Wales lose population. The nine other non-metropolitan zones continue to grow, particularly the Outer South East (18%), East Anglia (11%) and the South West (7%). All eight metropolitan zones lose shares of the national population together with the Northern Remainder. The non-metropolitan zones are gainers, particularly those already mentioned together with the East Midlands and the West Midlands Remainder.

8.3 The effect of closure choices recognizing external migration

From the tables, particularly Table 13, we can see that models 1, 5, 6 and 7 give very similar results. These are the models using the same accounting method (forecast model), which recognize external migration but employ slightly different techniques for handling those flows. Models 1 and 5 are virtually identical in their results - the differences are probably only due to slight computational differences. Whether you use net migration rates, gross migration rates, gross migration flows does not appear, at least in the short run, to have a profound effect on the projection outcomes.

8.4 The effect of ignoring external migration

The effect of ignoring external migration is quite substantial. Model 4 differs from the norm, Model 7, by three quarters of a million in the total UK population and has a substantial sum of absolute differences over the zonal populations, and the sum of absolute percentage differences, which gives each region in the analysis an equal weight, is high.

8.5 The effect of using a world system

The result of opening up the UK to world influences is, as expected, substantial in that overall growth increases and a rise of just over 800 thousand is projected by 2006 in the national population, as a result of the increased levels of immigration generated by a fast growing world population.

8.6 The effect of accounting model choice

When we depart from the simple forecast model view that all the accounts components can be accepted at face value in the forecast model, and use instead a backcast or constrained model substantial changes are effected in the population of the UK and paricularly in its distribution among the zones. The indices of dissimilarity, which compare the zonal distributions of each model with the norm, are high for Models 2 and 3 (Table 13). In the backcast model all the accounting errors are loaded on the initial populations; in the constrained model the emigration values are the ones adjusted to take the strain. Both adjustments in the view of what is currently happening in the base period result in level changes in zonal populations in 2006 of the same order as those resulting from adopting a world system model (see columns 1 and 2 of Table 13), and distributional changes that are larger (see columns 3 and 4).

A very careful examination of the base period accounts and likely error sources is thus indicated. It is clearly foolish to contemplate that the future can be accurately predicted if the past has been erroneously measured.

9. CONCLUSIONS

It is probably too early in the investigation of this new UK multi-regional system to recommend a replacement for the forecast accounting, emigration rates and immigration flows model (Model 7). More of the research design matrix (Table 7) needs to be explored, and the base period population changes need to be investigated thoroughly. However, it has been demonstrated that the two sources of variation in population projection outcomes that have been investigated in this paper, namely, accounting model choice and closure model choice, can be neglected by the forecaster only at his or her peril.

REFERENCES

Alexander, S. (1981) A model of population change with new and return migration. Environment and Planning A, 15, 1231-1257.

Central Policy Review Staff (1976) Population and the social services. Central Statistical Office, London.

Craig, J. (1983) The growth of the elderly population. Population Trends, 32, 28-33.

Easterlin, R. (1980) Birth and fortune. Grant McIntyre, London.

Martin and Voorhees Associates and John Bates Services (1981) Developing the migration component of official sub-national population projections. Final Report. Prepared for DPRP3 Division, Department of the Environment, London.

Ogilvy, A.A. (1979) Migration - the influence of economic change. Futures, 11 , 5, 383-394.

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

Rees, P. and Willekens, F. (1981) Data bases and accounting frameworks for IIASA's Comparative Migration and Settlement Study. Collaborative Paper, CP-81-39, International Institute for Applied Systems Analysis, Laxenburg, Austria.

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

Rogers, A. (1976) Shrinking large-scale population projection models. Environment and Planning A, 8, 515-541.

Rogers, A. (1983) Population projections. Chapter 9 in A. Rogers and F. Willekens (ed.) Migration and settlement:a comparative study. Academic Press, forthcoming.

Rogers, A. and Ledent, J. (1976) Increment-decrement life tables: a comment. Demography, 13, 287-290.

Stillwell, J.C.H. (1983) Analysis of internal migration in the UK based on transition and movement data. Paper presented at the Anglo-Dutch Migration Symposium, Soesterberg, Netherlands, 14-16 September 1983.

Stone, P.A. (1970) Urban development in Britain ... Vol.1. Population trends and housing. National Institute of Economic and Social Research, Economic and Social Studies, XXVI. Cambridge University Press, Cambridge.

Willekens, F., Por, A. and Raquillet, R. (1981) Entropy, multiproportional and quadratic techniques for inferring detailed migration patterns from aggregate data. IIASA Reports,4,83-124.

Willekens, F. and Rogers, A. (1978) Spatial population analysis: methods and computer programs. Research Report RR-78-18. International Institute for Applied Systems Analysis, Laxenburg, Austria.