Working Paper 268

Population geography: a review of model building efforts .

Philip Rees

Paper prepared for N. Wrigley and R. J. Bennett (ed) Quantitative Geography in Britain: Retrospect and Prospect

Routledge and Kegan Paul Ltd, Henley, Oxfordshire

School of Geography
University of Leeds
Leeds LS2 9JT

December 1979

Contents

Contents (continued)

Abstract

A series of questions concerning the behaviour of geographical populations is posed, and the paper attempts to review how population geographers and others, in Britain and elsewhere, have sought to answer those questions and with what effect.

List of tables

List of figures

1. Quantitative methods in population geography

In Britain a relatively small band of geographers are actively engaged in research into the ways populations behave in space. The Population Geography Study Group has been one of the smaller, though active, study groups of the Institute of British Geographers. Members figure prominently in the survey of current population research in the United Kingdom in Jackson (1979) and some nine out of thirty one organizations listed as conducting 'demographic/social policy' research are departments of Geography (pp. 39-41 in Jackson, 1979).

Almost all population geographers use numbers in their work, and so employ quantitative methods to analyse those numbers. If we adopt a rather tighter definition of quantitative methods as methods of analysis going beyond the simple construction of rates of incidence of population phenomena to a concern with how those rates of incidence can be explained quantitatively and with the consequences of the rates of incidence we find a much smaller number of active workers. If attention were to be confined to those studies which have gained attention in the wider field of demographic studies, and to those studies originating in Britain, my guess would be that we would have a rather small collection of work to examine (Figure 1).

This review of the questions posed and work carried out in quantitative population geography in Britain in the 1970s will therefore, of necessity, have to step outside the narrow confines of P∩Q∩B in Figure 1. The connections to a wider world must be made.

However, I am afraid the review will not satisfy many observers in the demographic studies set, D, nor the quantitative methods, set Q. Relatively little will be said, for example, about fertility patterns and trends, a dominant concern in demographic studies, because with a few exceptions (Compton, 1978; Jones, 1975) geographers have left this field to others. Quantitative geographers will also be rather dissatisfied with the way attention given to 'space' as a variable of import: space is treated in discrete chunks called regions in a rather practical fashion. Distance between regions does enter into migration models but not perhaps in an entirely satisfactory way.

What kind of questions have population geographers addressed? How have they sought to answer these questions and with what results? Of general and local British interest have been the following list of questions and associated problems.

(i) How do regional and local populations grow (or shrink) over time?

(ii) Given a proper understanding of the first question, how are regional and local populations likely to grow, in the future? How do we forecast the likely population changes in sub-national spatial units?

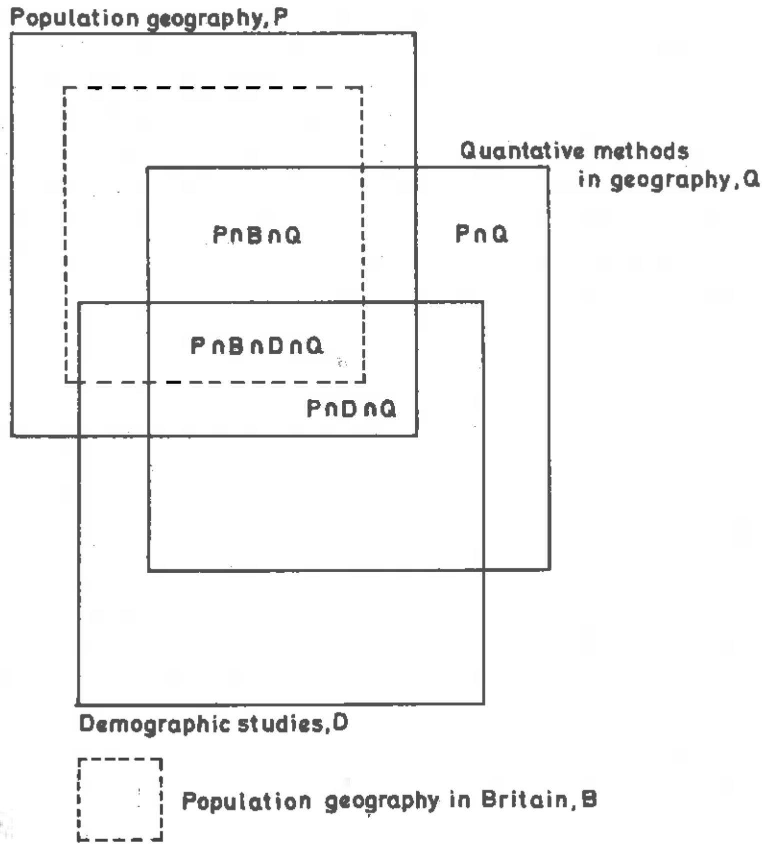

Figure 1. Some overlaps of topics and methods in population geography, quantitative methods and demographic studies

(iii) How can we investigate answers to the first two questions for many populations of complex character simultaneously? We need to study the evolution of multi-dimensional populations.

(iv) A number of questions need to be answered in the course of constructing models to handle the first three questions. How should our systems be closed? How should conflicting forecasts at different scales be resolved? What role should time (and time series) play in the analysis? Can we test our forecasting models?

(v) Recognizing that the movement of people over space plays a crucial role in regional and local population change, how can we predict the scale and direction of movement? Does our extensive experience of the modelling of migration really help in its prediction?

(vi) What influence does migration have on the life expectancies of people living in regions or local areas? What problems does the tendency of populations to remember their past have on models that tend to forget past history?

Subsequent sections of this review look at the answers suggested to these questions by population geographers and others, and an attempt is made to expose the problems that remain unsolved.

2. How do regional and local populations grow?

2.1 Simple components of growth

Several researchers (Department of the Environment, 1971; Lawton, 1977, 1980; Champion, 1976; Eversley, 1971; Stillwell, 1979) have examined in detail the pattern of population growth in British standard regions, subregions, or counties in the recent past, and others (Kennett 1977, 1980; Department of the Environment, 1976; Gleave and Cordey-Hayes, 1977) have described patterns for a system of city-regions. The equations underlying the work have been fairly simple, namely that

$$P^i(t+T) = P^i(t) + NI^i(t,t+T) + NM^i(t,t+T) \qquad (1)$$

where P^i refers to the population stock in region i, NI^i to the natural increase accruing to the population of region i and NM^i to the net migration experienced by the population in region i. The label (t) refers to the start of a period; (t+T) to a time T years later at the end of the period; and the two together, (t t+T), refer to the period as a whole. Subsequently, we will generally drop the period label (t,t+T) to make the equations easier to read.

Components of change can be defined by a simple manipulation of equation (1) into an equation identifying the sources of change in population numbers:

$$P^i(t+T) - P^i(t) = \Delta P^i = NI^i + NM^i \qquad (2)$$

where ΔP^i is the change in population are the period, and statistics such as

$$\frac{NI^i}{|NI^i + NM^i|} \quad \text{and} \quad \frac{NM^i}{|NI^i + NM^i|}$$

are used to measure the importance of natural increase and net migration respectively (Department of the Environment, 1971). Or, alternatively, the components of growth are converted into rates and the equation into one of rate of change:

$$\frac{\Delta P^i}{P^i(t)} = \frac{NI^i}{P^i(t)} + \frac{NM^i}{P^i(t)} \qquad (3)$$

or. $$g^i = ni^i + nm^i \qquad (4)$$

adopting lower case letters for the growth, natural increase and net migration rates.

So a simple population projection equation would then be

$$P^i(t+T) = (1 + g^i)P^i(t)$$

$$P^i(t+T) = (1+ni^i+nm^i)P^i(t)$$

$$P^i(t+T) = P^i(t) + ni^iP^i(t) + nm^iP^i(t)$$

or

$$P^i(t+T) = P^i(t)(1+b^i-d^i) + nm^iP^i(t) \quad (5)$$

An alternative population at risk to that used in equation (3), the initial stock, might be more appropriate in purely historical studies - namely, $P^i(t+\frac{1}{2})$ or $\frac{1}{2}(P^i(t) + P^i(t+T))$ - but any associated forecasting model would have to be iterative.

In the case where the average population in a period had been used the forecasting model would have to be

$$P^i(t+T) = P^i(t) + ni^i\tfrac{1}{2}(P^i(t) + P^i(t+T)) + nm^i\tfrac{1}{2}(P^i(t) + P^i(t+T)) \quad (6)$$

$$P^i(t+T) = P^i(t) \quad \frac{(1 + \frac{1}{2}ni^i + \frac{1}{2}nm^i)}{(1 - \frac{1}{2}ni^i - \frac{1}{2}nm^i)} \quad (7)$$

This need to match the projection model with the population at risk used in the input rate definition is frequently ignored in more complex contexts.

2.2 Empirical findings

The major empirical finding of the investigation of the components of growth in British city regions (Department of the Environment, 1976; Kennett, 1977, 1980) and standard regions (Rees, 1979a) was that net migration was the most influential component in determining the variation in population growth rate. The simple correlation between growth rate and natural increase rate for British standard regions for single years between 1965-66 and 1975-76 averaged +0.20 whereas the mean simple correlation of growth rate and net migration rate was +0.96 (Rees, 1979a, Table 6, p.24). Such results provide "moral" support for the geographer's view of the world against the overwhelming emphasis placed by demographers on the fertility component of natural increase, which at the national scale is much more important, of course.

2.3 A multiregional view of population change

It has long been recognised that the simple view of components of population growth covers a multitude of population movements, into and out of regions. The components of growth can be written in the following terms if gross migration flows are recognised (along with the birth and death components of natural increase):

$$P^i(t+T) = P^i(t) + (B^i \qquad - D^i \qquad) + (\sum_{j\neq i} M^{ji} - \sum_{j\neq i} M^{ij}) \tag{8}$$

where B^i and D^i refer to the births and deaths in region i and M^{ij} is the number of migrations between region i and region j. The internal migration term M^{ii} may be included in the equation but, if this is done, it will cancel out of course. This equation is converted into a projection equation by division of the flow terms by the origin zone population

$$P^i(t+T) = (1 + b^i - d^i - \sum_{j\neq i} m^{ij})P^i(t) + \sum_{j\neq i} (m^{ji}P^j(t)) \tag{9}$$

where $b^i=B^i/P^i(t)$, $d^i=D^i/P^i(t)$ and $m^{ij}=M^{ij}/P^i(t)$

Note that the in-migration to region i is a function of the out-migration rates from other regions multiplied by the populations of those regions.

Rogers (1966, 1968) has shown how the equation (9) model, can conveniently be set up as a matrix model.

If we let s^i be the rate at which persons survive in region i between time t and t+T, then

$$s^i = 1 - d^i - \sum_{j\neq i} m^{ij} \tag{10}$$

and the Rogers matrix model of population growth is (in a row vector form)

$$\left[P^1(t)P^2(t) \quad \dots \quad P^N(t) \right] \times \begin{bmatrix} s^1+b^1 & m^{12} & \dots & m^{1N} \\ m^{21} & s^2+b^2 & \dots & m^{2N} \\ \vdots & \vdots & \ddots & \vdots \\ m^{N1} & m^{N2} & \dots & s^N+b^N \end{bmatrix}$$

$$= \left[P^1(t+T)P^2(t+T) \quad \dots \quad P^N(t+T) \right] \qquad (11)$$

The Rogers model has been used by Compton (1969) in an analysis of the population dynamics of Hungary, and in modified form for a set of British regions (Rees, 1976, 1977a). Although the model has been superseded by more sophisticated life-table or accounts based versions it still constitutes a useful tool for learning about the dynamics of multi-regional growth (see Rogers, 1976 for examples of its use in combination with other models).

One of the major issues that ha exercised planners or civil servants concerned with population projection is whether a "net migration" perspective was adequate. In a survey of contemporary local authority practice in 86 authorities Woodhead (1979) found that 48% (Tables 3, 4) still employ a projection model incorporating net migration compared with 48% using gross flows (and 4% not taking migration into account at all!). The official sub-national projections (Campbell, 1976) still employ a methodology incorporating net migration although the Department of the Environment has recently let out to contract the design of a model for projection incorporating gross flows between areas (Girling, 1979). More will be said on this issue in the next section in connection with projection.

2.4 An accounting view of population change

The representation of population change embodied in equations (8), (9) and (11) involves sets of event counts (births, deaths) and moves counts (migrations). An alternative representation is to define a variable the superscripts of which define transitions of a population between states over a period. An accounts matrix can be defined as follows:

$$\underline{K} = \left|\begin{array}{cccc|cccc} K^{\varepsilon(1)\sigma(1)} & K^{\varepsilon(1)\sigma(2)} & \ldots & K^{\varepsilon(1)\sigma(N)} & K^{\varepsilon(1)\delta(1)} & K^{\varepsilon(1)\delta(2)} & \ldots & K^{\varepsilon(1)\delta(N)} \\ K^{\varepsilon(2)\sigma(1)} & K^{\varepsilon(2)\sigma(2)} & \ldots & K^{\varepsilon(2)\sigma(N)} & K^{\varepsilon(2)\delta(1)} & K^{\varepsilon(2)\delta(2)} & \ldots & K^{\varepsilon(2)\delta(N)} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ K^{\varepsilon(N)\sigma(1)} & K^{\varepsilon(N)\sigma(2)} & \ldots & K^{\varepsilon(N)\sigma(N)} & K^{\varepsilon(N)\delta(1)} & K^{\varepsilon(N)\delta(2)} & \ldots & K^{\varepsilon(N)\delta(N)} \\ \hline K^{\beta(1)\sigma(1)} & K^{\beta(1)\sigma(2)} & \ldots & K^{\beta(1)\sigma(N)} & K^{\beta(1)\delta(1)} & K^{\beta(1)\delta(2)} & \ldots & K^{\beta(1)\delta(N)} \\ K^{\beta(2)\sigma(1)} & K^{\beta(2)\sigma(2)} & \ldots & K^{\beta(2)\sigma(N)} & K^{\beta(2)\delta(1)} & K^{\beta(2)\delta(2)} & \ldots & K^{\beta(2)\delta(N)} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ K^{\beta(N)\sigma(1)} & K^{\beta(N)\sigma(2)} & \ldots & K^{\beta(N)\sigma(N)} & K^{\beta(N)\delta(1)} & K^{\beta(N)\delta(2)} & \ldots & K^{\beta(N)\delta(N)} \end{array}\right| \quad (12)$$

where ε and β are initial life states, of existence and of birth, in a period, and σ and δ are final states, of survival and of death, in a period. These lifestates have regional labels attached in brackets. The theory of population accounts originates with Stone (1965, 1971, 1972, 1975) and has been developed in a spatial unit context by Rees and Wilson (1973, 1975, 1977), Wilson and Rees (1974). and by Illingworth (1976). Examples of accounts for British regions are given in Rees (1976, 1977a and 1979a).

Row sums of this matrix may be defined as

$$K^{\varepsilon(i)*(*)} = \sum_j K^{\varepsilon(i)\sigma(j)} + \sum_j K^{\varepsilon(i)\delta(j)} \quad (13)$$

and

$$K^{\beta(i)*(*)} = \sum_j K^{\beta(i)\sigma(j)} + \sum_j K^{\beta(i)\delta(j)} \quad (14)$$

and similarly the column sums are

$$K^{*(*)\sigma(j)} = \sum_i K^{\varepsilon(i)\sigma(j)} + \sum_i K^{\beta(i)\sigma(j)} \quad (15)$$

and

$$K^{*(*)\delta(j)} = \sum_i K^{\varepsilon(i)\delta(j)} + \sum_i K^{\beta(i)\delta(j)} \quad (16)$$

If the set of regions, 1, 2 ... N, are defined so as to close the system (usually this means region N is a "rest of the world" region) then these marginal sums are identical with the population stocks and event counts used earlier in the components-of-growth model

$$P^i(t) = K^{\varepsilon(i)*(*)} \qquad i=1, \ldots, N-1 \quad (17)$$

$$P^j(t+T) = K^{*(*)\sigma(j)} \qquad j=1, \ldots, N-1 \quad (18)$$

$$B^i = K^{\beta(i)*(*)} \qquad i=1, \ldots, N-1 \quad (19)$$

$$D^i = K^{*(*)\delta(j)} \qquad j=1, \ldots, N-1 \quad (20)$$

The totals for the Nth and 2Nth rows and columns are a little different since normally $K^{\varepsilon(N)\sigma(N)}$, $K^{\varepsilon(N)\delta(N)}$, $K^{\beta(N)\delta(N)}$, and $K^{\beta(N)\delta(N)}$ are all set to zero if N is the "rest of the world" region. The total $K^{\varepsilon(N)*(*)}$ is then the total immigrants into the system of interest who either survive or die there, $K^{\beta(N)*(*)}$ is the corresponding total of infant immigrants. The column total $K^{*(*)\delta(N)}$ is the sum of surviving emigrants from the system of interest to the rest of the world and $K^{*(*)\delta(N)}$ is the total of non-surviving emigrants.

Once an accounts matrix has been estimated (see Rees and Wilson, 1977; Rees, 1978a; Rees, 1975b and Illingworth, 1976 for details of methods), the matrix model of population growth (equation (11)) can then easily be derived (see Rees and Wilson, 1977, Chapter 6). What then, you may ask, is the advantage of first estimating an accounts matrix before beginning a projection exercise? The advantages are as follows:

(i) Construction of the accounts matrix forces the analyst to check and adjust his base period data and to iron out inconsistencies. Illingworth, 1976, Jenkins and Rees, 1977a and Rees, 1979b give details of methods and provide evidence of the importance of this step. A variety of balancing factor, biproportional and RAS methods (Macgill, 1977, Bacharach, 1970, Stone, 1976 and Byron, 1978) can be applied to improve estimates of the accounts matrix of regional population change.

(ii) The analyst is forced to close the system under study. Frequently, models with equations based on (11) are misspecified if they lack a rest of the world region.

(iii) The analyst can use a variety of different models for projecting the different components of growth embodied in the accounts. He or she is not necessarily tied to the linear model of equation (11) (see Rees and Wilson, 1977, Chapter 7; Jenkins and Rees, 1977 and Rees, 1979b for various alternatives).

The disadvantages from which accounting has been said to suffer (Baxter and Williams, 1978; Cordey-Hayes, 1975) are several.

(i) A great deal of work of estimation is involved in preparing the data for input to the model that estimates the accounts.

(ii) The framework is purely involved with data assembly - no insights into the behaviour of the population are provided.

Both criticisms have validity but the conclusions which I would draw from them are rather different from those of the authors cited above. That a great deal of estimation work is necessary means that official agencies such as the Office of Population Censuses and Survey need to improve the organisation and reporting of their data bases. Population accounts constitute a plan for so doing.

The second criticism can be regarded as a strength. No particular model of population behaviour nor hypothesis about trends is imposed. But they can be tested, historically, against the data assembled in accounts.

2.5 An empirical example

Population accounts are, however, more than just data tables. They juxtapose population flows that would otherwise never be compared and force the analyst to estimate flows that would not otherwise be considered. Consider Table 1, for example, which shows a set of population accounts for the South East region for 1971-76.

From the table we can extract numerical examples of all the views of population change so far presented.

Population change for the South East, ΔP^{SE}, is negative

$$\Delta P^{SE} = \underset{\substack{\text{Population}\\ \text{at c.d.}\\ 1976}}{16{,}898.7} - \underset{\substack{\text{Population}\\ \text{at c.d.}\\ 1971}}{16{,}988.1} = -89.4 \qquad (21)$$

Table 1. A population accounts table for the South East region, 1971-76

Initial State \ Final State		Survivors at c.d. 1976			Deaths 1971-76			Totals
		South East	Rest of G.B.	Rest of World	South East	Rest of G.B.	Rest of World	
Existence at c.d. 1971	South East	14,670.3	734.5	612.0	941.6	13.3	16.4	16,988.1
	Rest of G.B.	587.5	33,607.7	627.9	6.8	2,202.6	18.7	37,051.2
	Rest of World	527.4	566.3	0	2.8	3.7	0	1,100.1
Births 1971-76	South East	1,049.4	34.8	32.2	19.4	0.3	1.4	1,137.5
	Rest of G.B.	27.0	2,489.6	33.6	0.2	48.2	1.6	2,600.2
	Rest of World	37.1	56.3	0	0.3	0.5	0	94.2
Totals		16,898.7	37,489.2	1,305.7	971.1	2,268.6	38.1	58,971.4

Notes: Source: Rees (1977a, Table 5, p. 15). c.d. = census date, April 24/25

1971-76 refers to period from c.d. 1971 to c.d. 1976

Figures are in 1000s. They have been slightly adjusted to ensure correct row and column summation.

(all figures in 1000s), and this is a product of positive natural increase and negative net migration:

=	(1,137.5	−	971.1)	+	(−255.8)
	Births 1971-76		Deaths 1971-76		Net migration 1971-76

$$= (166.4) + (-255.8) = -89.4 \qquad (23)$$

The net migration is the balance of two migration streams:

Total migrants to the South East	= (587.5)	+	(527.4)	exist-survive migrants	
	+ (27.0)	+	(37.1)	born-survive migrants	
	+ (6.8)	+	(2.8)	exist-die immigrants	
	+ (0.2)	+	(0.3)	born-die migrants	
	migrants from the Rest of Britain		migrants from the rest of the world		
	= 1189.1				(24)

Total migrants from the South East	= (734.5)	+	(612.0)	exist-survive migrants	
	+ (34.8)	+	(32.2)	born-survive migrants	
	+ (13.3)	+	(16.4)	exist-die migrants	
	+ (0.3)	+	(1.4)	born-die migrants	
	migrants from the Rest of Britain		migrants from the Rest of the World		
	= 1444.9				(25)

so that net migration into the South East is

$$NM^{SE} = 1189.1 - 1444.9 = -255.8 \qquad (26)$$

The accounts framework has forced us to recognise three types of migrants not normally considered – "born-survive" or surviving infant migrants, "exist-die" migrants or or non-surviving migrants, and "born-die" migrants or non-surviving infant migrants. Only "exist-survive" migrants are tabulated in census migration tables. We have also been obliged to estimate the number of emigrants from the South East who move abroad (the second column of numbers in equation (25)). External migration flows turn out to be very important in the population change equation for the South East. If we decompose the migrant streams into internal and external flows:

$$
\begin{aligned}
{}_{NM}^{SE} &= {}_{NM}^{SE\ \text{internal}} + {}_{NM}^{SE\ \text{external}} \\
&= (621.5 - 782.9) + (567.6 - 662.0) \\
&= (-161.4) + (-94.4) \\
&= -255.8 \qquad (27)
\end{aligned}
$$

it is apparent that external migrant flows are almost as large a component of change as internal migrant flows.

From the accounts a very general form of the Rogers' matrix growth model can be specified from the matrix of transition rates formed by dividing each row element by its row total (as in Table 2). The infant survival and migration rates can be multiplied by the birth rates and added to the exist-survive transition rates to give a growth rates matrix for 1971-76 (Table 3). This can then be used in a population growth model:

$$
\begin{aligned}
&[16{,}988.1 \quad 37{,}051.2] \begin{bmatrix} .9254 & .0453 \\ .0166 & .9743 \end{bmatrix} \\
&= [15{,}720.8 + 615.0 \quad 769.6 + 36{,}099.0] \\
&= [16{,}335.8 \quad 36{,}868.6] \qquad (28)
\end{aligned}
$$

to which must be added a vector of surviving immigrants and infant immigrants from abroad to yield the end of period population observed in the accounts*:

$$
\begin{aligned}
&[16{,}335.8 \quad 36{,}868.6] + [564.5 \quad 622.6] \\
&= [16{,}900.3 \quad 37{,}491.2] \qquad (29)
\end{aligned}
$$

The slight difference between the Table 1 and equation (29) populations is due to working with rates to only 4 decimal places. This causes slight rounding error.

2.6 The linkage of natural increase and net migration

So far, we have viewed natural increase and net migration as separate components. Of course, they are linked in that migrant flows will have associated with them natural increases (some migrants will die, and some will give birth to children). If the net balance of migrant flows is into a region, natural increase

*We might alternatively add rates of immigration to the diagonal elements of the growth rates matrix.

Table 2. Transition rate for the South East population accounts, 1971-76

Initial State \ Final State		Survive at c.d. 1976			Die 1971-76			Totals
		SE	RB	RW	SE	RB	RW	
Exist at c.d. 1971	SE	.8636	.0432	.0360	.0554	.0008	.0010	1.0000
	RB	.0159	.9071	.0169	.0002	.0594	.0005	1.0000
	RW	.5340	.5147	0	.0025	.0034	0	1.0000
Born 1971-76	SE	.9225	.0306	.0283	.0171	.0003	.0012	1.0000
	RB	.0104	.9574	.0129	.0001	.0185	.0006	1.0000
	RW	.3938	.5977	0	.0032	.0053	0	1.0000

Table 3. Steps in the construction of the growth rates matrix for the South East, 1971-76

Birth rates

$$\begin{array}{c} \text{SE} \\ \text{RB} \end{array} \begin{bmatrix} .0670 \\ .0702 \end{bmatrix}$$

Birth and survival rates

$$\begin{array}{ccc} & \text{SE} & \text{RB} \\ \begin{array}{c} \text{SE} \\ \text{RB} \end{array} & \multicolumn{2}{c}{\begin{bmatrix} .0618 & .0021 \\ .0007 & .0672 \end{bmatrix}} \end{array}$$

Growth rates matrix

$$\begin{array}{ccc} & \text{SE} & \text{RB} \\ \begin{array}{c} \text{SE} \\ \text{RB} \end{array} & \multicolumn{2}{c}{\begin{bmatrix} .8636 + .0618 & .0432 + .0021 \\ .0159 + .0007 & .9071 + .0672 \end{bmatrix}} \end{array}$$

or

$$\begin{array}{c} \text{SE} \\ \text{RB} \end{array} \begin{bmatrix} .9254 & .0453 \\ .0166 & .9743 \end{bmatrix}$$

Growth rates matrix with immigration rates

$$\begin{bmatrix} .9254 + .0332 & .0453 \\ .0166 & .9743 + .0168 \end{bmatrix}$$

=

$$\begin{bmatrix} .9586 & .0453 \\ .0166 & .9911 \end{bmatrix}$$

will be higher than it would otherwise have been, and it will be lower if net migration is balance out of a region. This linkage can clearly be demonstrated over several time periods in a projection model, but can the effect be detected within one period? Using accounts one can detect this effect.

Let us consider three types of population - the stayers, the outmigrants and the immigrants. Births and deaths occur to these three different kinds of persons. The births total for a region is made up of births to stayers, births to out-migrants (before outmigration) and births to immigrants (after in-migration). The deaths total is made of deaths to stayers plus deaths to immigrants. For the South East we might estimate these terms as

$$\begin{bmatrix} b^{*SE}{}_K\varepsilon(SE)\sigma(SE) \\ +\tfrac{1}{2}b^{*SE}{}_K\varepsilon(SE)\sigma(R) \\ +\tfrac{1}{2}b^{*SE}{}_K\varepsilon(R)\sigma(SE) \end{bmatrix} = \begin{bmatrix} 1{,}049.5 \\ +48.2 \\ +39.9 \end{bmatrix} = 1{,}137.6$$

in the case of births (R is the sum of regions other than the South East) where

$$b^{SE} = {}_K\beta(SE)^{*}(*) / {}_K\hat{\beta}^{*SE} = .071536$$

$${}_K\hat{\beta}^{*SE} = {}_K\varepsilon(SE)\sigma(SE) + \tfrac{1}{2}{}_K\varepsilon(SE)\sigma(R) + \tfrac{1}{2}{}_K\varepsilon(R)\sigma(SE)$$

$$= 15{,}901.5$$

The deaths terms for stayers are

$${}_K\varepsilon(SE)\delta(SE) + {}_K\beta(SE)\delta(SE) = 941.6 + 19.4 = 960.0$$

the out-migrant deaths are

$${}_K\varepsilon(SE)\delta(R) + {}_K\beta(SE)\delta(R) = 29.7 + 1.7 = 31.4$$

and the in-migrant deaths are

$${}_K\varepsilon(R)\delta(SE) + {}_K\beta(R)\delta(SE) = 9.6 + 0.5 = 10.1$$

Only the first and last terms sum to the deaths total, however.

It could be said that the process of migration meant that the South East had (48.2 - 39.9) or 8.3 thousand fewer births than if no one had migrated. If no one had migrated the South East would have experienced 21.3 thousand more deaths. The effect of net out-migration on natural increase was therefore to increase it by -8.3 + 21.3 = 13.0 thousand from 166.4 to 179.4 thousands.

This is not the only way in which the interactions between natural increase and net migration could be traced, and the effects are not as might first be supposed. Of course, a wide variety of other factors - the age structure of the population and migrants, the specific fertility schedules of the various migrant streams, the differential mortalities of migrant and stayer streams - affect the outcome, and a sophisticated projection model is probably needed to unravel all the effects.

2.7 Views of population change: prospect

What are the problems that remain for our views of population change, and how likely is that descriptions of the population geography of Britain which will use the results of the 1981 census will adopt the multi-regional or accounting framework rather than the simple components-of-growth?

There are still a number of theoretical problems needing resolution. These concern the ways in which judgement about the statistical reliability, the variances of each of the sets of terms that enter accounts-population estimates (some error variance), population census counts (less error variance), migrant counts (sampling error, misallocation error), births and deaths data (pretty small errors) - can be incorporated in the methods of estimation. Current practice is outlined in Rees (1979b) but this has been rightly criticized by Keyfitz (1979). The solution probably lies in the adaptation of methods suggested some time ago by Stone, Champernowne and Meade (1942) and elaborated more recently by Stone (1975b) and Byron (1978). A best estimate of accounts matrix is made, taking into account the sizes of the error variances about the estimates of the elements of the accounts. The problem still remains, however, of making judgemental estimates of those error variances.

There are also implementation problems. Computer programs for developing multi-regional accounts need improvement before they can be widely used, particularly in more complex age-sex disaggregated form. The data series also need upgrading, particularly those for migration. Census type retrospective migrant data are the correct data for use in accounts (see Rees, 1977b) but in Britain we get this data only with the census. It would be very simple, cheap and effective for the Office of Population Censuses and Surveys to give some spatial detail in their General Household Survey reports (O.P.C.S., Social Survey Division, 1973, 1974, 1975, 1976, 1977). The purist might say that the sample sizes for any elements in a migration table would yield standard errors of the estimate of rates or proportions that would be too broad. But any numbers would be useful in this context since in between censuses all that the researcher has to go on are indirect partial and unsatisfactory sources such as the National Health Service Register moves (to be used in the Department of Environment project, Girling, 1979) or statistics on electoral roll numbers or housing starts and completions.

However, it is interesting to note that a full registration system such as that operated in the Netherlands, Denmark, Sweden, Norway, Finland, Czechoslovakia, Hungary, and other countries does not actually yield the number of transitions (migrants) between two regions over a period but rather the number of moves (see Ledent, 1978a, 1978b, 1979 for a discussion of the consequences of using one set of data or another). Transitions data is the data required in multi-regional population models of the Rogers type, or in models derived from population accounts. From the point of view of stock projection use of movements data makes little difference: equation (8) still holds and moves "surplus" to transitions (see Illingworth, 1976 and Rees, 1977b for a discussion of the concept of "surplus" moves) cancel out. However, any statistics based on the probability of making a transition between states which use movement data must be in error because

$$p^{ij}(t,t+T) \neq M^{ij}(t,t+T)/P^{i}(t) \qquad (30a)$$

Instead

$$p^{ij}(t,t+T) = K^{\varepsilon(i)\sigma(j)}(t,t+T)/K^{\varepsilon(i)*(*)} \qquad (30b)$$

where $p^{ij}(t,t+T)$ is the probability of being in state j at time t+T given that you were in state i at time t.

Whether the multi-regional or accounting frameworks will be adopted for analysis of the 1981 census data really depends on the models adopted for forecasting the population of spatial units. To these attention is now turned.

3. How are regional and local populations likely to grow?

3.1 Preliminaries

It is planners (national, regional and local) who have been most concerned with forecasting population rather than geographers, so we must step outside the P set into the D set in Figure 1 if we are to understand the issues involved. The terms 'forecasting' and 'projection' are used fairly interchangeably here as endeavours to guess what will happen in the future, though they have often been given separate meanings (Pittenger, 1976; Brass, 1979) of 'best guess' and 'conditional prediction' respectively.

3.2 Multiregional models versus single region models with net migration

The principal arguments for the adoption of multiregional methods have been outlined by Rogers and Philipov (1979). The most important may be put as follows. Starting from the same base period data and making the same assumption that the migrant pattern remains constant, we would observe migration flows in the multiregional case responding to the fluctuating size of origin region population stocks whereas in the net migration case there would still be net migrants even if the populations of other regions disappeared. This effect is even more serious if net flows are used rather than net rates.

This point can be illustrated using our South East example. From the accounts table (Table 1) and associated transition rates table (Table 2) can be computed the rates needed for both net migration based and multiregional models. Two versions of the net migration based model are used, and these are called model (1) and model (2).

Model (1), single region model with net migration flow

Here we use rates to model the natural increase component and flows to represent the net migration (as in the official sub-national projections):

$$P^i(t+5) = P^i(t)(1 + b^i - d^i) + NM^i \tag{31}$$

where for the base period, 1971-76, and subsequently,

$$\begin{array}{llll} b^{SE} = .0670 & & b^{RB} = .0702 \\ d^{SE} = .0572 & & d^{RB} = .0612 \\ NM^{SE} = -255.8 & & NM^{RB} = 106.4 \end{array}$$

Model (2), single version model with net migration rates

If we model net migration in rate terms we obtain the second model (as in equation (5) earlier)

$$P^i(t+T) = P^i(t)(1 + b^i - d^i + nm^i) \tag{32}$$

with, for 1971-76 and subsequently

$$nm^{SE} = -.0151 \qquad nm^{HB} = .0029$$

so that

$$g^{SE} = -.0053 \qquad g^{HB} = .0119$$

and

$$1+g^{SE} = .9947 \qquad 1+g^{HB} = 1.0119$$

Model (3), multiregional model with immigrant flow

This is the model we used informally in equation (28)

$$\underline{p}(t)\ \underline{H} + \underline{m} = \underline{p}\ (t+5) \tag{33}$$

where $\underline{p}$ is a row vector of regional populations, $\underline{H}$ a matrix of inter-regional transition rates plus birth rates (multiplied by infant transition rates), and $\underline{m}$ is a vector of surviving immigrants from abroad (plus surviving infant immigrants). The structure of $\underline{H}$ is as follows:

$$\underline{H} = \begin{bmatrix} (s^{11} + b^1\ is^{11}) & (m^{12} + b^1\ im^{12}) & \dots & (m^{1N} + b^1\ im^{1N}) \\ (m^{21} + b^2\ im^{21}) & (s^{22} + b^2 is^{22}) & \dots & (m^{2N} + b^2\ im^{2N}) \\ \vdots & \vdots & & \vdots \\ (m^{N1} + b^N\ im^{N1}) & (m^{N2} + b^N\ im^{N2}) & \dots & (s^{NN} + b^N\ is^{NN}) \end{bmatrix} \tag{34}$$

where is^{ii} refer to the rate of infant survival within a region and im^{ij} the rate of infant migration from region i to region j. In the South East example, we use

$$\underline{H} = \begin{bmatrix} .9254 & .0453 \\ .0166 & .9743 \end{bmatrix} \tag{35}$$

$$\underline{m} = \begin{bmatrix} 564.5 & 622.6 \end{bmatrix} \tag{36}$$

Only survivors are represented explicitly in this equation, and the net migration figures generated are not comparable with those of the single region model, so it is sometimes advisable to decompose and expand the model to deal with all the terms in the accounts

$$\underline{p}(t)\ \underline{H}_{ES} + \underline{m}_{ES} = \underline{p}_{ES}(t) \tag{37}$$

$$\underline{p}(t)\ \underline{H}_{ED} + \underline{m}_{ED} = \underline{p}_{ED}(t) \tag{38}$$

$$\underline{p}(t)\,\underline{B}\,\underline{H}_{BS} + \underline{m}_{BS} = P_{BS}(t) \tag{39}$$

$$\underline{p}(t)\,\underline{B}\,\underline{H}_{BD} + \underline{m}_{BD} = P_{BD}(t) \tag{40}$$

where E refers to existence, S to survival, D to death and B to birth.

Model (4), multiregional model with immigrant rate

The same matrix equation is used as in Model (3), except that immigration rates are added to the diagonal terms in the $\underline{H}$ matrices and the $\underline{m}$ vectors are dropped

$$\underline{p}(t)\ \underline{H} = p\,(t + 5) \tag{41}$$

where the diagonal terms in $\underline{H}$ are now the sum of 4 rates:

$$(s^{ii} + m^{Ri}) + (b^{i}\ is^{ii} + im^{Ri})$$

The $\underline{H}$ we use is, computed from the accounts,

$$\underline{H} = \begin{bmatrix} .9586 & .0453 \\ .0166 & .9911 \end{bmatrix} \tag{42}$$

The results of the application of models (1) through (4) are summarized in Table 4 through to the year 2041 (the 1971-2041 span representing the life expectancy of the author's son). Population, share and net migration statistics are presented for 1971 (base point in time), 2001 and 2041.

The differences in behaviour of the two regional populations are remarkable; none of these differences can be attributed to data differences nor to assumptions (all models assume the 1971-76 situation persists). There are small deviations to rounding error differences in the arithmetic (of $\pm$ 1 thousand in the 1971-76 period in the net migration part of Table 4), but the major discrepancies can be attributed to model form.

The graphed results (Figures 2 through 5) show the differences most clearly. In the single region model the populations of the South East and the Rest of Great Britain continue on the paths mapped out in 1971-76 - decline for the South East and growth for the rest of Britain. The second model shows simple compound growth or decay; in the first model this is modified by a constant item that has decreasing impact on the population trajectory of the growing region but increasing impact on the trajectory of the declining region.

In contrast, the growth paths of the regions in the multiregional model show convergence. The South East ceases to decline in 2021-26 in model (3) and 2026-31 in model (4). Net migration loss becomes substantially less for the

Table 4. Summary of the results of the projections of models (1) through (4)

Year	Single region models				Multiregional models			
	Model (1)		Model (2)		Model (3)		Model (4)	
	POPULATION: SOUTH EAST (SHARES)							
1971	16,988.1	(31.44)	16,988.1	(31.44)	16,988.1	(31.44)	16,988.1	(31.44)
2001	16,438.9	(29.26)	16,455.0	(29.26)	16,645.2	(29.66)	16,611.7	(29.56)
2041	15,654.8	(26.43)	15,770.1	(26.51)	16,647.1	(28.33)	16,566.6	(27.97)
	POPULATION: REST OF GREAT BRITAIN (SHARES)							
1971	37,051.2	(68.56)	37,051.2	(68.56)	37,051.2	(68.56)	37,051.2	(68.56)
2001	39,750.5	(70.74)	39,776.6	(70.74)	39,480.8	(70.34)	39,582.1	(70.44)
2041	43,582.8	(73.57)	43,724.9	(73.49)	42,119.4	(71.67)	42,666.4	(72.03)
	POPULATION: GREAT BRITAIN (SHARES)							
1971	54,039.3	(100.00)	54,039.3	(100.00)	54,039.3	(100.00)	54,039.3	(100.00)
2001	56,189.4	(100.00)	56,231.6	(100.00)	56,126.0	(100.00)	56,193.8	(100.00)
2041	59,237.6	(100.00)	59,495.0	(100.00)	58,766.5	(100.00)	59,233.0	(100.00)
	NET MIGRATION: SOUTH EAST							
1971-6	- 255.8		-256.5		- 253.9		- 254.1	
2001-6	- 255.8		-248.5		- 192.8		- 201.0	
2036-41	- 255.8		-239.4		-143.4		- 143.4	
	NET MIGRATION: REST OF GREAT BRITAIN							
1971-6	106.4		107.4		105.8		105.2	
2001-6	106.4		114.0		18.9		50.7	
2036-41	106.4		125.3		-78.6		-11.0	
	NET MIGRATION: GREAT BRITAIN							
1971-6	-149.4		-149.1		-148.1		-148.9	
2001-6	-149.4		-134.5		-173.9		-150.3	
2036-41	-149.4		-114.1		-222.0		-154.5	

Note

The population figures are in 1000's

The share figures are percentages

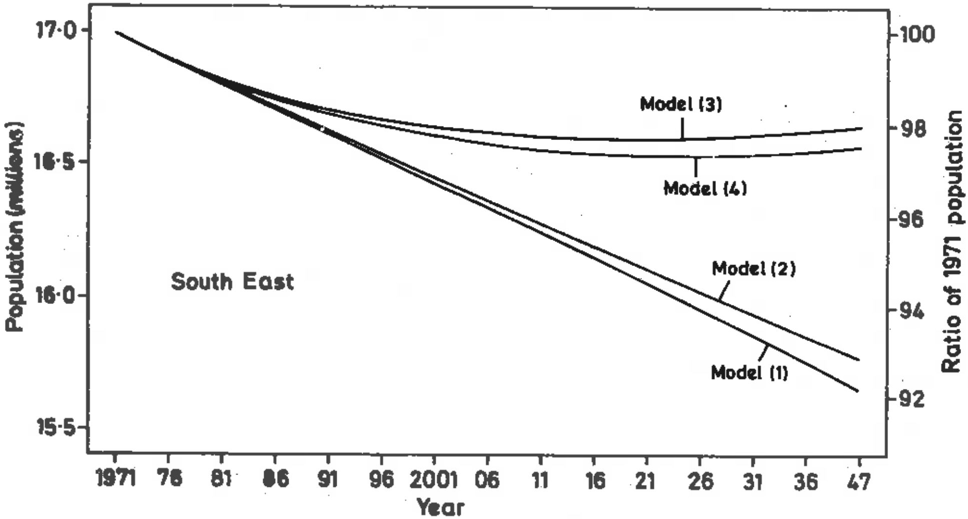

Figure 2. Projected populations of the South East under models (1) through (4)

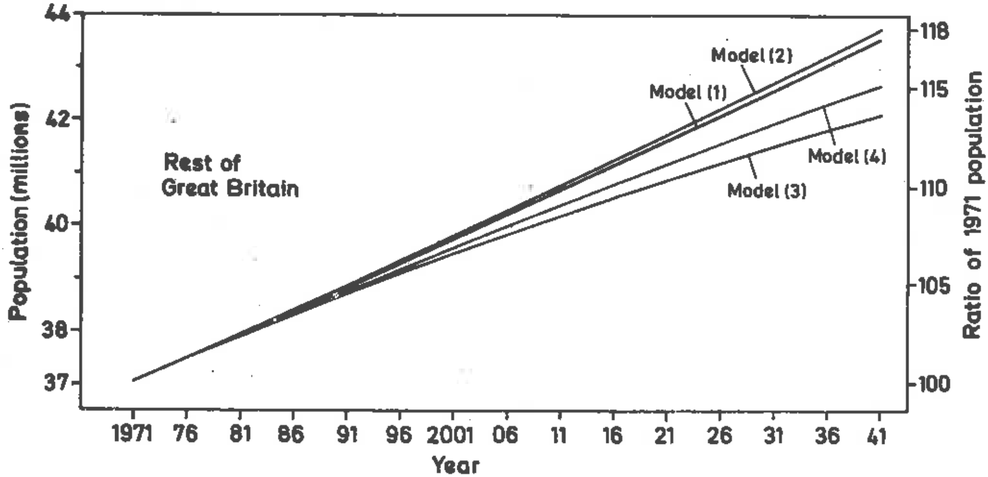

Figure 3. Projected populations of the Rest of Great Britain under models (1) through (4)

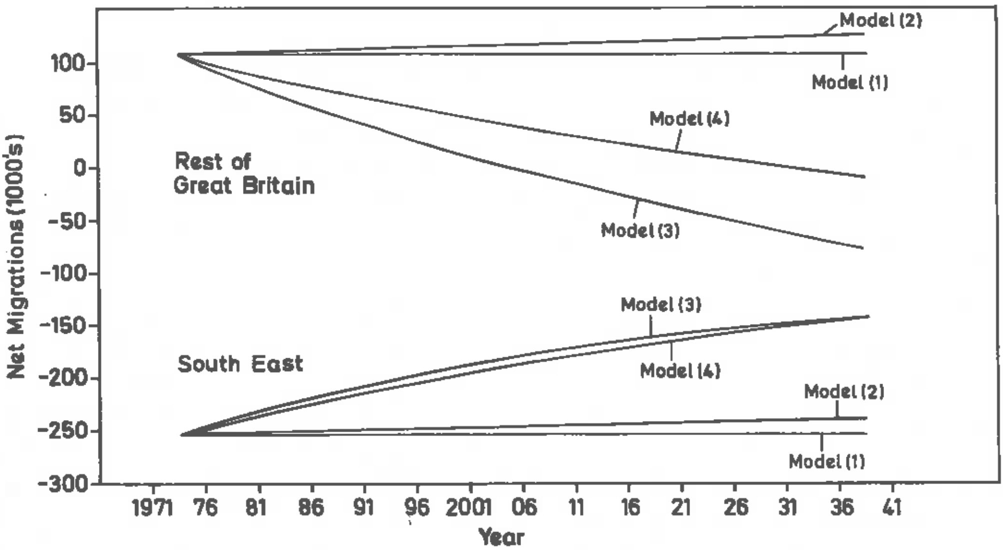

Figure 4. Projected net migrations of the South East and the Rest of Great Britain under models (1) through (4)

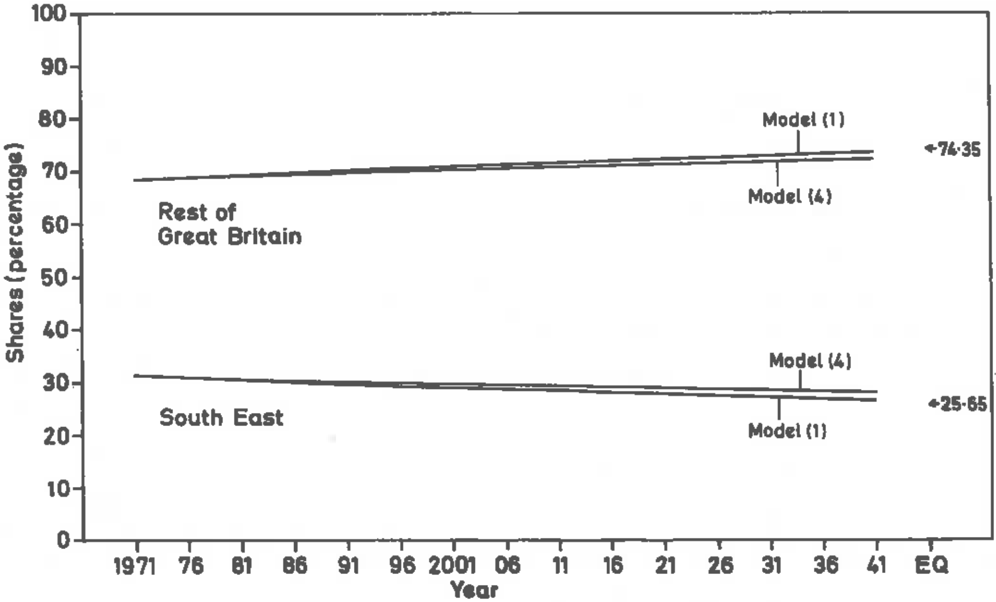

Figure 5. Projected shares of the Great Britain population under models(1) through (4)

South East and net migration gain falls for the Rest of Great Britain until losses are recorded towards the end of period. These shifts in pattern are due, of course, to feedback effects: as the Rest of Great Britain grows so does the size of the migration stream to the South East; as the South East declines so does the size of the migration stream to the Rest of Great Britain. Consequently, the net migration losses of the South East to the Rest of Great Britain are reduced.

It is often interesting to investigate what the ultimate consequences of a long term application of a projection model might be in terms of the long run growth rates of the regions and their shares of the national population. For models (1) and (2) ultimately the population of the South East reaches zero and that of the Rest of Great Britain continues to grow indefinitely at the rate observed in 1971-76. For model (4), Rogers (1968, Chapter 5) showed that the dominant root of the characteristic equation associated with the growth matrix was the intrinsic rate of growth of the population in the system at stability. For the growth matrix in this case we have

$$\det \begin{vmatrix} .9586-\lambda & .0453 \\ .0166 & .9911-\lambda \end{vmatrix} = 0 \tag{43}$$

where det $\begin{vmatrix} \quad \end{vmatrix}$ refers to the determinant of matrix, which is

$$(.9586-\lambda)\ (.9911-\lambda) - (.0453)\ (.0116) = 0 \tag{44}$$

$$\lambda^2 - 1.9497\lambda + .94931648 = 0 \tag{45}$$

whose roots are $\lambda_1 = 1.006725422$

or $\lambda_2 = 0.9429745784$

The populations of both the South East and Rest of Britain ultimately grow at a rate of 1.0067 per quinquennium or 6.7 per 1000.

The population shares at equilibrium are found from the equations

$$\underline{p}\ \underline{H} = \underline{p}\lambda \tag{46}$$

$$\underline{p}\ \underline{H} - \underline{p}\lambda = 0 \tag{47}$$

$$\underline{p}\ (\underline{H} - \lambda\underline{I}) = 0 \tag{48}$$

Numerically, in the case of example these are

$$(p^1\ p^2) \left\{ \begin{pmatrix} .9586 & .0453 \\ .0166 & .9911 \end{pmatrix} - \begin{pmatrix} 1.006725422 & \\ 0 & 1.006725422 \end{pmatrix} \right\} = (00) \tag{49}$$

from which an expression for p_1 in terms of p_2 can be found

$$p^1 = \{.0166/(.9586 - 1.006725422)\} \; p_2 \tag{50}$$

$$p^1 = .3449 \; p^2 \tag{51}$$

then if we set $p^1 + p^2 = 1$

we obtain $p^2 = 1/1.3443 = .7435$ (52)

$$p^1 = 1 - .7435 = .2565$$

Model (3) with its constant vector will vary from this asympt tic behaviour, but show stable behaviour when the constant terms are reduced to very minor additions to the populations of the regions.

Which of the models should we choose to project the regional populations? The multiregional models exhibit more reasonable behaviour in the long run. Whether model (3) or model (4) is chosen depends on one's view of immigration from abroad. This is a term subject in part to legislative control and there might be a case for representing explicitly as a flow vector rather than a rate. The internal migration terms, in Britain, are not subject to such legislative control and there is no corresponding case for representing them as flows.

3.3 Multiregional matrix models versus multiregional accounting models for forecasting

The numerical terms input to our 'toy' models for the South East and the Rest of Britain were all derived from the Table 2 accounts, but they could have been estimated, in principle, directly from the various data sources. Let us assume we could calculate birth rates, death rates, internal migration rates, and net external migration rates from commonly available British data:

	South East	Rest of Great Britain
b^i	.0670	.0702
d^i	.0572	.0612
m^{ij}	.0432	.0159
nm^{Ri}	-.0062	-.0015

Then the terms in the $\underline{H}$ matrix would be estimated as follows:

$$h^{ii} = (1 - d^i - \sum_{j \in I} m^{ij} + nm^{Ri}) \, (1 + b^i) \tag{53}$$

$$h^{ij} = m^{ij} + b^i \, (m^{ij}) \tag{54}$$

with the result

$$\underline{\mathbf{H}} = \begin{bmatrix} .9533 & .0461 \\ .0170 & .9877 \end{bmatrix} \qquad (55)$$

Compared with equation (42)'s estimation, the diagonal terms are underestimated and the off-diagonal over-estimated. A careful analysis of the accounting meaning of the terms in equations (53) and (54) would reveal considerable misspecification.

The degree of misspecification of the input variables varies from study to study. Liaw's projections of the Canadian multi-provincial system (Liaw, 1977, 1978a, 1978b, 1979) are very well specified, for example, although they were not developed from an accounting base. Other studies, however, neglect minor migrant flows or external migration, although the authors usually claim that the omissions are not important (Gleave and Cordey-Hayes, 1977, p. 29; Baxter and Williams, 1978, p. 43).

However, another and stronger argument can be put forward for first calculating an account matrix before projecting the population. This is that in estimating accounts the consistency and reliability of the input data are assessed, and often the initial accounts matrix is adjusted. These arguments are set out in detail in Rees (1979b), but an illustration is appropriate here.

The migration flows over the one year before the census between the North Yorkshire and Humberside and North West regions are given in the 1971 Census as (when multiplied by 10)

From: (April 24/25, 1970)	To: (April 24/25, 1971) North	Yorks & Humberside	North West
North	-	9,820	6,740
Yorks & Humberside	8,760	-	12,630
North West	6,900	9,470	-

After the accounts have been constructed in an unconstrained version for the census year 1970-71 and subsequently constrained to the population estimated at mid-year 1970 and 1971 using balancing factor or RAS methods (Jenkins and Rees, 1977) the resulting estimates are (Rees, 1979b, p. 28):

From: June 30, 1970	To: June 30, 1971		
	North	Yorks & Humberside	North West
North	-	10,145	6,820
Yorks and Humberside	8,508	-	12,397
North West	6,840	9,687	-

The North to Yorkshire and Humberside figure, for example, changes from 9,820 to 10,145 as a result of adjusting the accounts matrix, although the revised figure still lies within the 95% confidence limits for this flow of 9,210 to 10,430 . The migration rate shifts from .0031 to .0032 as a result.

4. How can we handle many populations of complex character simultaneously?

4.1 General issues

Any demographer or population geographer worth his salt would, of course, throw up his hands at the projection models presented earlier in this section. We have to disaggregate the population by age and sex, and this is done for the preferred multiregional models by the authors involved (Rogers, 1968, 1975; Willekens and Rogers, 1978; Rees and Wilson, 1977; Plessis-Fraissard and Rees, 1976) in both theoretical and applied contexts. The details of the models involved are given in these references. Here we consider the problems posed.

There are problems involved in such disaggregation, however. The number of variables involved and the size of the matrixes employed are increased. If five year age groups are used the number of variables is increased by say 18 x 2 or 36. If one year age groups are used then the number of variables goes up by say 85 x 2 or 170. There may be 170 separate versions of half of Table 1 (since this table already involves two 'age groups') as a result, for example. This may be acceptable for a table containing 2 $(N^2 - 1)$ elements when N is 3 making 170 x 16 or 2720 variables in total, but when N is 108 (total number of non-metropolitan counties (39), metropolitan districts (36) and London boroughs (33) in England and Wales) explicit representation of all elements $(170 \times 2 (108^2 - 1) = 3,965,420)$ will result in an unreliably estimated matrix with a majority of elements being zero and the average being about 15.

Several approaches to this dimensionality problem have been tried. The first is the development of aggregated models and an investigation of their performance in comparison with disaggregated models. The second is the use of chained probability equations, often in a hierarchic fashion. Both approaches involve information loss but gain through reduction in the number of elements involved.

4.2 Experiments with aggregated and decomposed models

Rogers (1976) has experimented with a wide variety of aggregated and decomposed models for a system consisting of the populations of the 9 Census divisions of the United States with 1955-60 as the base period. Projections of the population were carried out for a full 9 by 9 multiregional model and then various aggregated and decomposed alternatives were compared with the full 9 by 9 model. Table 5 assembles a set of goodness of fit statistics from the results reported by Rogers. The goodness of fit statistics compare the ability of the various models to match the regional populations (total) projection at 2008 by the full multiregional models.

In first model in Table 5 Rogers aggregates the 9 x 9 Census division system

Table 5. Comparisons of different aggregations/decompositions with the full 9 x 9 multiregional model at 2008

Model Number	Model description	Difference in US total population	Goodness of fit statistics Absolute value of differences	Sum of the absolute percentage differences	Index of dissimilarity
(1)	4 region model	-397,125	1,854,003	1.91	0.22
(2)	Components of growth model	57,253,458	57,253,458	112.07	1.59
(3)	Bi-regional aggregation	-6,675,293	7,698,812	16.01	0.74
(4)	Single region with net mig	-28,919,917	54,700,122	108.74	6.83
(5)	Decomposition A with net mig	-19,589,551	37,060,124	77,32	4.76
(6)	Decomposition B with net mig	-13,372,645	25,665,116	53.37	3.28
(7)	Single region with components of growth	2,688,865	13,636,737	28.55	1.59
(8)	Decomposition B with bi-regional aggregation	-1,307,861	5,431,940	10.94	0.62

Source: Computed from Rogers (1976)

Notes on the goodness of fit statistics

1. Difference in US total population for model n = K_9^{US} (2008) − K_n^{US} (2008) where K_n^{US} refers to the total population of the US as projected by model number n and n = 9 refers to the full 9 by 9 multiregional model. Positive values indicate underprediction and negative values over prediction vis à vis the full multiregional model. K_9^{US} (2008) = 421.862 millions.

2. Absolute value of differences for model n = $\sum_{i=1}^{9} \left| K_9^i (2008) - K_n^i (2008) \right|$ where K_n^i is the total population of census division (or region) i projected under model n.

3. Sum of the absolute percentage differences for model n = $\sum_{i=1}^{9} 100 \times$

$$\left| \frac{K_9^i (2008) - K_n^i (2008)}{K_9^i (2008)} \right|$$

. This gives each regional projected population an equal weighting.

4. <u>Index of dissimilarity for model n</u> $= \frac{1}{2}\, 100 \sum_{i=1}^{9} \left| \frac{K_9^i\ (2008)}{K_9^{US}\ (2008)} - \frac{K_n^i\ (2008)}{K_n^{US}(2008)} \right|$

This compares the relative distribution of the populations among divisions in two models.

to a 4 x 4 Census region system with relatively good results for those 4 aggregated units. The problem, of course, is that this doesn't provide predictions of the 9 Census division populations. The second model, a simple components of growth model (with no age groups distinguished), underpredicts, as one might expect, but does relatively well in projecting the regional shares, as the index of dissimilarity (a relative distribution statistic) shows.

The third model involves aggregation of the 9 x 9 model into 9 sets of 2 x 2 models. This model performs much better than the second, because, of course, age information is introduced, and it also outperforms the next three models that involve tearing the 9 by 9 systems apart in various ways and compensating for this with net migration terms. The seventh model in which a components of growth model is combined with a single region, natural increase-only model appears to be a compromise that works relatively well although it may not produce regional age breakdowns that properly reflect the influence of migration. The final model involves tearing of the 9 by 9 system into 3 sets of 3 by 3 systems compensated by aggregating the other regions into a 'rest of the country' category. Although this last model does better than either the bi-regional aggregation model and single region with components of growth models, the gain does have its price. Model (8) demands migration data tables in the same degree of detail as the full 9 by 9 model. Model (7) gets by with an aggregate multiregional migration table only, and model (3) gets by with total in- and out-migration disaggregated by age. These latter two data tables are much more likely to be available than full age-sex disaggregated multiregional data tables, especially when there are large numbers of regions (zones). Another advantage of models (3) and (2) over those involving decomposition ((5), (6) and (8)) is that they are general procedures, applicable to any multiregional system, irrespective of the pattern of interregional flows. Design of the best decomposition involves careful investigation of the particular flow system and design of a model and computer program appropriate to that system.

4.3 A migration accounting framework

However, when modelling population systems with large numbers of regions this careful investigation and specific model building is clearly worthwhile. Masser (1976) outlines a hierarchical accounting system that could well be adapted for large multiregional projection models. Figure 6 illustrates the way in which the Masser's accounting system works for a ten zone system.

The small units, the migration patterns of which are being investigated or for which projections are to be carried out, are grouped into blocks of regions. Within each block all interregional flows are accounted for explicitly. The

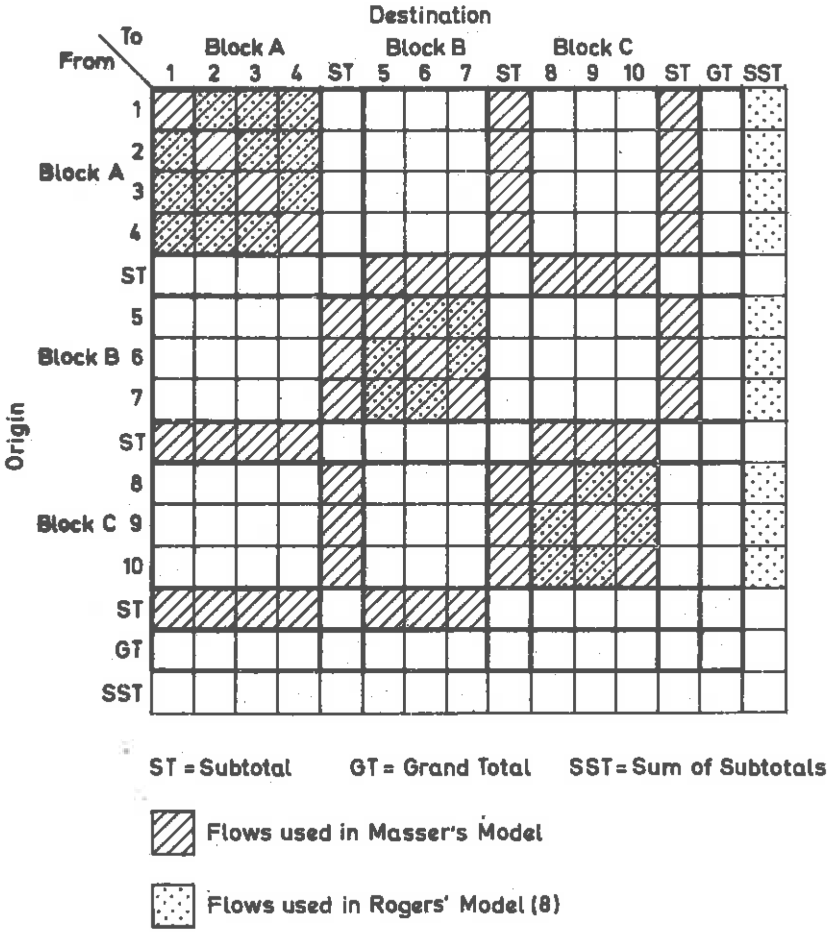

Figure 6. An accounting framework for a decomposed two-level spatial system (Adapted from Masser, 1976, Figure 1 and Rogers, 1976, Figure 5)

flows from regions in one block to regions in another are only accounted for in terms of subtotals, but consistency is maintained:

$$M^{i_I OUT} = \sum_{j_I} M^{i_I j_I} + \sum_{k \neq I} M^{i_I K} \tag{56}$$

$$M^{IN j_J} = \sum_{i_J} M^{i_J j_I} + \sum_{K \neq J} M^{K j_J} \tag{57}$$

where i_I represents the label for a region i in the Ith block of regions. The out-migration total, $M^{i_I OUT}$, is made of the sum of outflows to regions j within block I, and the sum of outflows to blocks K. Similarly, the in-migration total $M^{IN j_J}$ is made of the sum of inflows from other regions i within block J, and of the sum of inflows from other blocks K.

Model (8) in Rogers (1976) turns out to be a slightly greater aggregation than this in which the migration to the rest of the system, say $M^{i_I R}$, is substituted for all the $M^{i_I K}$ in Masser's scheme. The terms used in Rogers' model (8) are line shaded and the $M^{i_I R}$ are represented in the sum of subtotals columns.

The meaning and role of the diagonal elements in Figure 6's accounting framework can vary. There are some four possibilities.

(1) The entries in the $M^{i_I i_I}$ cells are migrations within the region. If the migration pattern is being modelled using a spatial interaction model, Masser recommends the flows be included as model calibrations are rather sensitive to their inclusion/exclusion. These migrations will cancel out in later population accounting equations.

(2) Often, as in the case with Masser's Dutch data, the intra-regional flows are missing and the migration flows must be modelled without these terms.

(3) From a population accounting point of view the proper entry in these diagonal cells should be a count of surviving stayers. This may be derived from census migration tables via the end of period population accounting equation.

(4) Or, as in many of the earlier models, the diagonal terms may be estimated as a residual (one minus the death rate minus the outmigration rate).

What savings are made by the aggregations carried out in Masser's accounting framework? These savings will be important, even if fully explicit data exist, in reducing the zero elements in the migration matrix, making it less sparse and therefore more reliable from a statistical point of view. The savings will also make the job of translating migration or population models into operational computer programs easier.

Let us assume we have N regions or smallest data base units (DBUs) as Masser calls them, and L blocks of regions (aggregates of DBUs). In Figure 6 N equals 10 and L equals 3. The blocks are labelled $l = 1, 2, \ldots, L$, and the number of regions in a block l is N_l, so that

$$\sum_{l=1}^{L} N_l = N \tag{58}$$

The number of elements in a fully explicit representation is

$$N \times N = N^2$$

whereas the number of elements in a Masser type framework is

$$M = \sum_{l=1}^{L} (N_l^2) + 2 \sum_{l=1}^{L} (N_l(L - 1)) \tag{59}$$

The savings, S, are therefore $S = N^2 - M$ or in proportional terms $\%S = \frac{N - M}{N} 100$. The first term on the RH side of equation (59) gives the number of intra-block interactions to be considered, and the second set of terms represent the off-block-diagonal elements to be represented. For Figure 6,

$$N^2 = 100$$
$$M = (4^2 + 3^2 + 3^2) + 2(4 + 3 + 3)(3 - 1)$$
$$S = 100 - 74 = 36, \%S = 36$$

The general savings can be worked out if we simplify the definition of M by assuming that blocks have equal numbers of units in them (on average) so that

$$M = L\left(\frac{N}{L}\right)^2 + 2(L - 1)N \tag{60}$$

If we wanted to study the metropolitan economic labour area (MELA) system of Britain defined in Department of the Environment (1976) and used the standard regions as the areas within which blocks of MELAs were found, then

$$N = 126, L = 10$$
$$N^2 = 15{,}876$$
$$M = 10\left(\frac{126}{10}\right)^2 + 2(9)126$$
$$= 3{,}856 \text{ (to the nearest integer)}$$
$$S = 15{,}876 - 3{,}856 = 12{,}020$$
$$\%S = 76$$

The number of elements would be reduced to 24% of those required in an explicit multiregional representation. Clearly, the larger N is the greater the savings achieved by using Masser's accounting representation. The savings also depend on the number of blocks, L; when there are few blocks there are few savings, and similarly when L is close to N there are few savings also. In between there will be a value of L with a minimum value of M associated with it.

Regarding N as a constant we may write

$$M = f(L) = L\left(\frac{N}{L}\right)^2 + 2(L-1)N$$

$$= N^2 L^{-1} + 2LN - 2N \qquad (61)$$

The first derivative is

$$f'(L) = -N^2 L^{-2} + 2N \qquad (62)$$

Setting this to zero

$$-N^2 L^{-2} + 2N = 0 \qquad (63)$$

we obtain an expression for L with given N,

$$L = \left(\frac{N^2}{2N}\right)^{\frac{1}{2}} \qquad (64)$$

which, since the second derivative is positive,

$$f''(L) = 2N^2 L^{-3} \qquad (65)$$

means that the associated M value is a minimum. In the study of 126 MELAs, the L value which gave the minimum M would be $\left(\frac{126^2}{2 \times 126}\right)^{\frac{1}{2}}$ or 8, and the minimum M value would be $8\left(\frac{126}{8}\right)^2 + 2(7)(126) = 3{,}749$.

4.4 Choice of blocks of regions

Choice of the number and composition of blocks would not solely be determined by the savings in elements, of course. Masser and Brown (1975) have discussed the adaptation of Ward's (1963) grouping procedure for flow systems. The distance measure adopted is the difference between observed and expected flows:

$$x_{ij} = (a_{ij} - a^*_{ij}) + (a_{ji} - a^*_{ji}), \; i \neq j \qquad (66)$$

where $a_{ij} = \dfrac{M^{ij}}{M^{**}}$, $M^{**} = \sum_i \sum_j M^{ij}$

and $a^*_{ij} = a_{i*} \cdot a_{*j}$

The M^{ij}s are replaced by proportions a_{ij} to prevent the grouping procedure from being dominated by large flows to a 'nodal' region. A contiguity constraint is also applied.

Two criteria can be used in the grouping procedure. Either the proportion of the total interaction that takes place within the blocks can be maximized or it can be minimized. In the case discussed above maximum intrablock interaction is the desired objective as it is these flows that will be represented in most spatial detail (by origin and destination region). Frequently, however, other criteria, external to the migration analysis intervene in the choice of block, or initial regionalization of the data base units. Population projections are normally prepared for administrative or planning purposes, and their boundaries may or may not be optimal from a migration analysis point of view.

4.5 Probability chain models

The issues of dimensionality and aggregation had to be faced by the Population Studies Section of the Greater London Council (Gilje and Campbell, 1973). Required was a model that would handle population projections for the 33 boroughs, 4 external zones (the outer Metropolitan Area, the Outer South East, the Rest of Great Britain and the Rest of the World), disaggregated by single years of age and by sex.

The solution involved grouping the boroughs into blocks in a fashion similar to that of Masser (as in Figure 7) together with a set of 4 external zones. The model uses multiregional 'rates' within Greater London but models the interaction to and from external zones in the form of ex geneously specified flows, apart from flows to the outer Metropolitan Area and Outer South East which are modelled in the same way as flows between boroughs in different blocks (Figure 7). The migrant flow between boroughs is modelled not by a migration rate multiplied by a population at risk (say, the initial population) but by two chains of more aggregated probabilities, one chain applying to intra-block flows, another to inter-block flows.

The intra-block equations* used by Gilje and Campbell can be represented as a probability estimation equation and as migrant flow estimation equation in which the estimated probability is multiplied by a population at risk:

$$\hat{p}\,(l,\ x/k,\ x-1) = p\,(k,\ l/m,\ A).\ p\,(m,\ x/A,\ x-1) \qquad k, l \in A \tag{67}$$

and

$$\hat{M}^{kl}_{x-1\,x} = \hat{p}\,(l,\ x/k,\ x-1)\ P^{k}_{x-1}(t) \qquad k, l \in A \tag{68}$$

where

$$p\,(m,\ x/A,\ x-1) = \frac{\sum\limits_{\substack{k\in A\\ k\neq l}} \sum\limits_{\substack{l\in A\\ l\neq k}} M^{kl}_{x-1x}}{\sum\limits_{k\in A} P^{k}_{x-1}(t)} \tag{69}$$

$$p\,(k,\ l/m,\ A) = \frac{\sum\limits_{x} M^{kl}_{x-1x}}{\sum\limits_{\substack{k\in A\\ k\neq l}} \sum\limits_{\substack{l\in A\\ l\neq k}} \sum\limits_{x} M^{kl}_{x-1x}} \tag{70}$$

---

* This notation differs radically from that in the Gilje and Campbell (1973) paper but is used to analyze the nature of their model more fully.

Destinations / Origins		A Central group 1 ... 4	B Inner group 5 ... 14	C Outer group 15 ... 33	External zones OMA OJE RGB	RW
Central A group	1 City of London . . . . 4 Westminster	Intra-group equations	Inter-group equations	Inter-group equations	Inter-group equations	Total migration flow allocated
Inner B group	5 Hackney . . . . 14 Tower Hamlets	Inter-group equations	Intra-group equations	Inter-group equations	Inter-group equations	Total migration flow allocated
Center group	15 Barking . . . . 33 Waltham Forest	Inter-group equations	Inter-group equations	Intra-group equations	Inter-group equations	Total migration flow allocated
The Outer Metropolitan Area OMA The Outer South East OSE The Rest of Great Britain RGB The Rest of the World RW		Total migration flow allocated	Total migration flow allocated	Total migration flow allocated	Not considered	

Figure 7. The spatial structure of the Greater London boroughs' model of Gilje and Campbell (1973)

The probability $\hat{p}$ (1, x/k, x - 1) is the predicted probability that a person living in borough k in (single year) age group x - 1 at the start of the year will survive in borough 1 in age group x at the end of the year. This is used in equation (68) to predict the migrant flow, $M^{k\,l}_{x-1\,x}$ between boroughs k and 1 of persons making the x - 1 to x age transition.

This expected probability is a product of two observed probabilities. The right most term in equation (67) is the probability that a person aged x - 1 living in a borough in borough group A will migrate to another borough in group A (this is the meaning of 'm') and survive in age group x. Equation (69) shows how this would be measured from observed data. The second observed probability is the probability that a person will migrate between borough k and borough 1 given that he or she has migrated from one borough to another within borough group A. Note that the origin-destination pair selection here is assumed to be independent of age.

The inter-block equations have a related structure

$$\hat{p}\,(1,\ x/k,\ x-1) = p\,(k,\ 1/A,\ B)\ .\ p\,(B,\ x/A,\ x-1)\,, \qquad k\in A,\ l\in B \tag{71}$$

and

$$\hat{M}^{k\,1}_{x-1x} = \hat{p}\,(1,\ x/k,\ x-1)\ P^{k}_{x-1}\,(t)\,, \qquad k\in A,\ l\in B \tag{72}$$

where

$$p\,(B,\ x/A,\ x-1) = \frac{\sum_{k\in A}\sum_{1\in B} M^{k\,1}_{x-1\,x}}{\sum_{k\in A} P^{k}_{x-1}\,(t)} \tag{73}$$

$$p\,(k,\ 1/A,\ B) = \frac{\sum_{x} M^{k\,1}_{x-1\,x}}{\sum_{k\in A}\sum_{1\in B}\sum_{x} M^{k\,1}_{x-1\,x}} \tag{74}$$

The definitions of the terms in the inter-block equations parallel those of the intra-block equations very closely. For the migrations between boroughs and the Outer Metropolitan Area (OMA) and Outer South East (OSE) the inter-block or group equations were used with the destination block, B, being the OMA or OSE zone. The other external migration terms were modelled as flows

$$\sum_{e} M^{e\,1}_{x-1\,x} = p\,(x-1,\ x/A)\ I^{1},\ 1\in A \tag{75}$$

where e refers to an external zone, I^{1} to a predetermined total in-migration flow to borough 1 from external zones and p (x - 1, x/A) is probability of such a migrant into boroughs in group A being in age transition x - 1 to x, where

$$p\,(x-1,\ x/A) = \frac{\sum_{e}\sum_{1\in A} M^{e\,1}_{x-1\,x}}{\sum_{e}\sum_{1\in A}\sum_{x} M^{e\,1}_{x-1\,x}} \tag{76}$$

This treatment of external flows makes the Greater London model only partially a multiregional model since the interaction between Greater London and the rest of the country is restricted.

It is instructive to now ask: what gains have been made in representing the multiregional transition probabilities as a chain in the equations (67) and (71)?

The terms $\hat{p}$ (1, x/k, x - 1) will contain $\sum_B N_B^2$ x R elements in the intra-block equations, when N_B is the number of boroughs in block B and R is the number of age groups (R^2 are not needed as people move only from one age group to the next). The terms p(k, 1/m, A) will contain $\sum_B N_B^2$ elements and p (m, x/A, x - 1) will contain L x R elements. The model thus gains in parsimony since $(\sum_B N_B^2) + LR < \sum_B N_B^2 \times R$.

A careful examination of the observed probability definitions indicates that variables, as detailed as $M^{kl}_{x\,x}$, are used to define the probabilities, so that real gains in parsimony occur only if the sum variables involved are available as independent totals rather than as sums of the detailed inter-borough, age to age flows. Given the way the borough blocks are defined, these independent totals are unlikely to be available (and the same points apply also to Masser's accounting framework). Thus, we could define

$$p(1, x/k, x-1) = M^{kl}_{x-1x} / P^{k}_{x-1}(t) \tag{77}$$

directly as an observed probability and avoid the loss of information in substituting $\hat{p}$ (1, x/k, x - 1) for p (1, x/k, x - 1).

The problem with this direct approach (in effect, adopting a straight-forward multiregional cohort survival model for systems with large numbers of regions) is that of 'matrix sparseness' and of the unreliability of the estimates of individual elements. Many of the terms in large matrices may be zero or very small, and if we regard these as samples (as the migration elements often are) then the standard errors are likely to be relatively high. In the long run it may be better to measure probabilities from more aggregate data and combine the probabilities in probability chain equations.

The penalty that has to be paid is that although many conditional probabilities are included in the chain, at some point the hypothesis of independence between two characteristics of the population must be invoked. Otherwise there would be no modelling gain. What is not done in the Gilje and Campbell study is the testing of the goodness of fit of the resulting model. No idea is given of the sacrifices in reproducing the observed data that have been made to gain an increase in reliability.

One final observation should be made here. Although the Masser two-level spatial system and Gilje and Campbell two-tier spatial hierarchy are very similar in concept they are put to rather different uses. Masser deliberately uses the hierarchy to reduce the number of elements to be modelled. Gilje and Campbell use the hierarchy to design two different models for intra- and inter-block flows but still produce estimates of the flows between every lower level unit.

4.6 Migration estimation models

The problem that Willekens and his colleagues (Willekens, 1977 and Willekens, Por and Raquillet, 1979) faced was to estimate the age disaggregated interregional migration matrix for countries in which this data was only partially available. The fully disaggregated data was required for input into a sophisticated multiregional population model program (Willekens and Rogers, 1978) as part of a comparative national study in IIASA's Migration and settlement task (Rogers, 1976b; Willekens, 1978) in order to study the dynamics of the regional population systems of seventeen countries (see Rees, 1979a and Rikhinen, 1979 for typical products).

A variety of different data situations and estimation methods are described in the Willekens (1977) and Willekens, Por and Raquillet (1979) papers. Here one situation and one method are described to provide the reader with a more concrete notion of what is involved. The variable that we are attempting to estimate is

$m_{ij}(x)$ = migrants between region i and region j in age group transition x to x + 1 (as usual we assume regular age groups and a simple xis diagram structure requiring only one age dimension)

in the form of a three dimensional array

$$\{ m_{ij}(x) \}$$

Known, and used as constraints in the estimation procedure are

$m_{ij}(\cdot)$ = the total number of migrants between region i and region j (the dot . indicates that the replaced index has been summed over)

$m_{.j}(x)$ = the total number of in-migrants to region j by age group x.

The arrays $\{m_{ij}(\cdot)\}$, $\{m_{i.}(x)\}$ and $\{m_{.j}(x)\}$ can be represented as matrices which fit together as faces of the cube formed by the three dimensional array $m_{ij}(x)$. Figure 8 is an attempt to show these relationships diagramatically.

Using the following definition of entropy

$$S = -\sum_i \sum_j \sum_x \left[m_{ij}(x) \ln m_{ij}(x) - m_{ij}(x) \right] \qquad (78)$$

the expression for $m_{ij}(x)$ that gives rise to the maximum (most probable) value of S, subject to

$$\sum_j m_{ij}(x) = m_{i.}(x) \qquad (79)$$

$$\sum_i m_{ij}(x) = m_{.j}(x) \qquad (80)$$

$$\sum_x m_{ij}(x) = m_{ij}(.) \qquad (81)$$

is

$$m_{ij}(x) = A_i(x)\, B_n(x)\, C_{ij}\, m_{i.}(x)\, m_{.j}(x) \qquad (82)$$

where $A_i(x)$, $B_j(x)$ and C_{ij} are parameters or balancing factors related to the Lagrange multipliers. Expressions for these are derived by substituting the RH side of equation (82) in equations (79), (80) and (81) respectively

$$A_i(x) = m_{i.}(x) / \left[\sum_j B_j(x)\, C_{ij}\, m_{i.}(x)\, m_{.j}(x) \right] \qquad (83)$$

$$B_j(x) = m_{.j}(x) / \left[\sum_i A_i(x)\, C_{ij}\, m_{i.}(x)\, m_{.j}(x) \right] \qquad (84)$$

$$C_{ij} = m_{ij}(.) / \left[\sum_x A_i(x)\, B_j(x)\, m_{i.}(x)\, m_{.j}(x) \right] \qquad (85)$$

Solutions to these equations must be derived iteratively. These techniques have been extensively used in input-output modelling (Stone, 1963; Bacharoch, 1970), in spatial interaction modelling (Wilson, 1970, 1974) and are reviewed comprehensively in Macgill (1975 and 1977, Chapter 3).

How are these estimations relevant to the problem of handling multi-dimensional population models with large numbers of elements? Assume that the $m_{ij}(x)$ are sample numbers only (as is the case in all UK census tabulations of migration to date). Then the standard errors of $m_{ij}(.)$, $m_{i.}(x)$ and $m_{.j}(x)$ will, relatively speaking, be much smaller than those of $m_{ij}(x)$. The standard errors of a predicted set of $m_{ij}(x)$'s should also be lower than the observed set. Extreme values in the observed $m_{ij}(x)$ set will be dampened in the predicted $m_{ij}(x)$ set - the data will have been smoothed. In the long term the values predicted by a model using aggregated data may be more 'reliable' than those using disaggregated data. These notions undoubtedly implicitly underpin the Gilje and Campbell model, and need careful further investigation

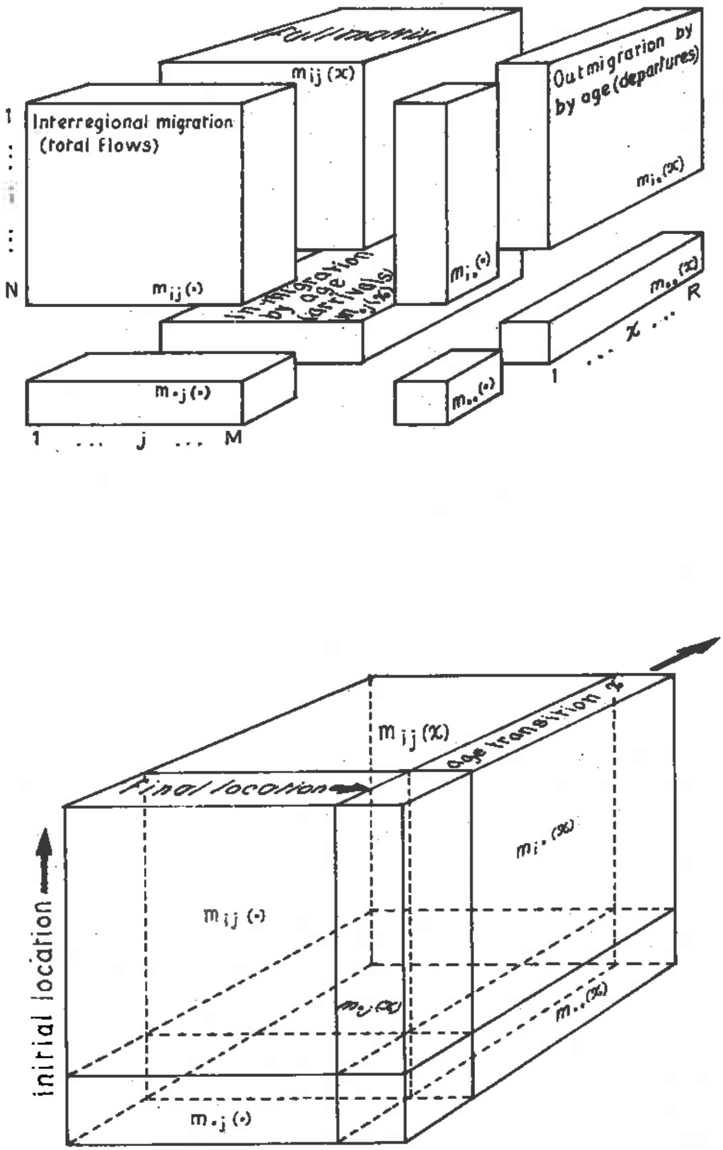

Figure 8. A diagramatic representation of the migration estimation problem (Adapted from Willekens, 1977, Figure 3)

in vigorous statistical argument and testing, particularly when the migration data derive from samples. A representation of the multiregional migration process in an equation akin to equation (82) does have the additional advantage that f r fewer input elements the need be projected. This has been realized for many years in the transport modelling field where large interaction problems a r e handled as a matter of course. The topic is taken up again in the discussion of migration models in section 6 of the review.

4.7 Migration age profiles

One of the reasons that multiregional models may run into dimensionality problems is that it is important to include age disaggregation in the models. The local planner undoubtedly would like this age disaggregation to as fine as single years where possible so that the $(2N)^2$ element of a multiregional accounts matrix or the $(N)^2$ elements of a multiregional population model must be multiplied by 85 or more.

Rogers, Raquillet and Castro (1978) have investigated ways in which the migration rate by age schedule may be modelled, thereby reducing the number of varying quantities in an associated population model. For R migration rates (where R may be 85+) are substituted n model parameters. They decompose the age specific migration rate, M(x), into four components (Figure 9).

$$M(x) = M_A(x) + M_B(x) + M_c(x) + M_D(x) \tag{86}$$

in which

$M_A(x)$ is the predicted pre-labour force migration rate component

$M_B(x)$ is the predicted labour force migration rate component

$M_c(x)$ is the retirement migration rate component

$M_D(x)$ is the base level migration rate.

The pre-labour force component is modelled as a simple negative exponential function of age:

$$\hat{M}_A(x) = a_1 e^{-\alpha_1 x} \tag{87}$$

The labour force component is modelled as a 'double exponential'

$$\hat{M}_B(x) = a_2 e^{-\alpha_2 (x - \mu_2) - e^{-\lambda_2 (x - \mu_2)}} \tag{88}$$

where μ_2 is a parameter indicating the age at which labour force migration begins to have an effect. The retirement component is modelled as a similar 'double exponential'

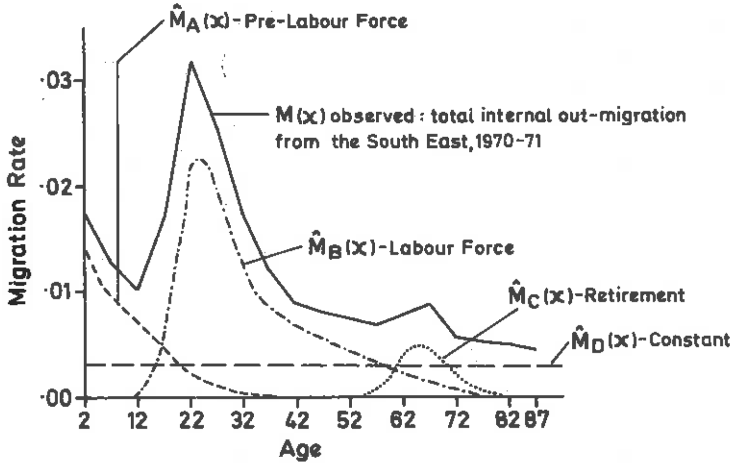

Figure 9. The fundamental components of the migration rate schedule. (Adapted from Rogers, Requillet and Castro, 1978, Figure 13A)

$$\hat{M}_c (x) = a_3 e^{-\alpha_3 (x - \mu_3)} - e^{-\lambda_3 (x - \mu_3)} \qquad (89)$$

where μ_3 is a parameter indicating the age at which retirement migration begins to have an influence. The final component is simply a base level constant.

$$\hat{M}_D (x) = C \qquad (90)$$

The full model therefore takes up the following form

$$\hat{M} (x) = a_1 e^{-\alpha_1 x} + a_2 e^{-\alpha_2 (x - \mu_2)} - e^{-\lambda_2 (x - \mu_2)} + a_3 e^{-\alpha_3 (x - \mu_3)} - e^{-\lambda_3 (x - \mu_3)} + C \qquad (91)$$

The full model has eleven parameters ($a_1, \alpha_1, a_2, \mu_2, \alpha_2, \lambda_2, a_3, \mu_3, \alpha_3, \lambda_3$ and C), and a reduced form without the retirement component would have seven. The 'savings' of this model representation would not be great for five year age group projection models but would effect significant 'savings' in the case of single year of age models. As in the previous section the model would tend to smooth the migration rate profiles and this might again be regarded as an advantage if the migration rates were based on samples. The gain in model element representation would, of course, be even greater if a national age profile model were combined with a gross interregional migration model (one of the estimation cases discussed by Willekens, 1977 and Willekens, Por and Raquillet, 1979).

5. Other issues in population projection

5.1 How should systems be closed?

One advantage of building a multiregional population model from a set of accounts which was identified earlier was that explicit consideration had to be given to proper closure of the system being studied. A variety of different options for model construction have been outlined by Jenkins and Rees (1977) and Rees (1979c). These are depicted in revised form in Figure 10.

The system being studied is divided into a set I of internal regions (within a country) and a set E of external regions. In the Table 1 example the South East and Rest of Britain are members of the internal set and the Rest of the World is the sole member of the external set. In Figure 10 it is assumed that the internal migrant terms in the model are represented as transition rates multiplied by a population at risk. Alternative 1 involves the use of net external migration flows to close the system; alternative 2 involves the use of net external migration rates, (as in Liaw, 1977). Alternative 3 leaves the analyst free to set emigrant and immigrant flow figures separately (as in Gilje and Campbell, 1973); alternative 4 models emigration using transition rates because the population at risk is known, and leaves the immigrant term to be determined exogeneously as a flow (as in Rees, 1977a). The flows may be modelled as a total flow into the internal system multiplied by a rate that distributes that total among the internal regions. Options 5 and 6 replace the flow terms in options 3 and 4 by admission rates in which the population at risk is the population of the destination region (model (4) in our earlier example was an option 6 model). Option 7 expands the internal system of interest to include the rest of the world explicitly.

From \ To	Internal set of regions, I	External set of regions, E
Internal set of regions, I	Internal stayers and migrants	Emigrants
External set of regions, E	Immigrants	Ignored

Alternative inputs

NF = net migration flow
NR = net migration rate x par
F = gross migration flow
TR = transition rate x par
FR = total immigrant flow x distribution rate
AR = admission rate x par
par = population at risk

1.

	I	E
I	TR	NF
E	NF	∅

2.

	I	E
I	TR	NR
E	NR	∅

3.

	I	E
I	TR	F
E	F/FR	∅

4.

	I	E
I	TR	TR
E	F/FR	∅

5.

	I	E
I	TR	F
E	AR	∅

6.

	I	E
I	TR	TR
	AR	∅

7.

	I	E
I	TR	TR
E	TR	TR

Figure 10. Options for closing the system

How much effect does choice of alternative means of closure of the system have? We have already seen that only fairly small differences were involved when option 4 (model (3))was chosen as opposed to option 6 (model (4)). Choosing option 7 makes a great deal of difference. Table 6 revises the Table 1 accounts by adding rest of the world population, births and deaths data; Table 7 extracts the transition rates from this table and Table 8 shows the results of applying 1971-76 rates through to 2041. As might be expected the Great Britain regions show much more growth than earlier because in this model migration from the Rest of the World is not restricted to a fixed flow nor by the origin region size but is simply dependent on a fast growing Rest of the World population. The South East, for example, ceases to decline after 1981, and is 14 per cent larger in size than under the model (3) and model (4) conditions reported earlier.

Which of the closure options should be adopted then? The option that best reflects the degree to which external migration is controlled must be adopted. Clearly, immigration into Britain is subject to many controls and is probably best represented as a flow, subject to the occasional surprise (as, for example, in 1972 when large numbers of Ugandan Asians were admitted). Although in the 1960's and earlier emigration from Britain was a fairly free process, restrictions have been introduced in the 1970's which have limited the numbers emigrating.

5.2 Top down or bottom up in forecasting?

It is usual for both central and local governments (and sometimes intermediate regional bodies) to carry out forecasts. Inevitably, these will differ because of differences in models and assumptions and the differences will be a matter of methodological and political debate. However, even if there were no such differences the central and local forecasts would still differ because of problems involved in aggregation. The populations projected

Table 6. <u>An extended accounts table for the South East region, 1971-76</u>

Initial state \ Final state		Survival at c.d. 1976			Death 1971-76			Totals
		SE	RB	RW	SE	RB	RW	
Existence at c.d. 1971	SE	14,670.3	734.5	612.0	941.6	13.3	16.4	16,988.1
	RB	587.5	33,607.7	627.9	6.8	2,202.6	18.7	37,051.2
	RW	527.4	566.3	3,426,793.2	2.8	3.7	212,350.2	3,640,243.6
Birth 1971-76	SE	1,049.4	34.8	32.2	19.4	0.3	1.4	1,137.5
	RB	27.0	2,489.6	33.6	0.2	48.2	1.6	2,600.2
	RW	37.1	56.3	538,740.1	0.3	0.5	16,188.3	555,022.6
Totals		16,898.7	37,489.2	3,966,839.0	971.1	2,268.6	228,576.6	4,253,043.2

<u>Notes</u>

1. The rest of the world population, births, deaths and stayer estimates are based on the figures in Brown (1976)

2. SE = South East, RB = Rest of Britain, RW = Rest of World

3. Figures are in 1000's.

Table 7. Transition rates associated with the extended accounts, 1971-76

Initial state \ Final state		Survival at C.D. 1976			Death 1971-76			Totals
		SE	RB	RW	SE	RB	RW	
Existence at c.d. 1971	SE	.863563	.043236	.036025	.055427	.000783	.000965	1.000000
	RB	.015856	.907061	.016947	.000184	.059447	.000505	1.000000
	RW	.000145	.000156	.941364	.000001	.000001	.058334	1.000000
Birth 1971-76	SE	.922549	.030593	.028308	.017055	.000264	.001231	1.000000
	RB	.010384	.957465	.012922	.000077	.018537	.000615	1.000000
	RW	.000067	.000101	.970663	.000001	.000001	.029167	1.000000

Table 8. The results of projections using the Tables 6 and 7 data

	South East	Rest of Britain	Rest of the World
Population (10000s)			
1971	16,898.1	37,051.2	3,640,243.6
2001	17,433.9	40,428.3	6,093,525.3
2041	22,257.6	49,775.1	12,101,204.1

for the whole system of interest for a more disaggregated system will always be greater than those projected by an aggregate model, (Keyfitz, 1977; Rogers and Philipov, 1979) unless the conditions for "perfect" aggregation are met (Rogers, 1969; Rogers 1971). For example, if the three region accounts of Table 1 are aggregated to a Great Britain and Rest of the World system and model (4) applied the 2041 population of Great Britain is 59.170 millions compared with 59.233 millions for the disaggregated system. Given this discrepancy, should forecasts for lower level regions be adjusted to those for higher level regions, or should the higher level region forecasts be simply an aggregation of the lower level? I am assuming that one body is carrying out the forecasts in this case. If both higher and lower level bodies are carrying out forecasts, the figures of the more important, higher level body usually win out.

The answer to the question probably depends on a judgement or investigation as to the reliability of lower level figures as opposed to higher. Thus, the question is a cousin to those raised at the end of section 4 and merits further research.

5.3 <u>What role should time series analysis play</u>?

So far we have assumed that population projections employ the base period rates <u>ad infinitum</u>. Clearly, this is not reasonable and most forecasters study the trends in key rates over the past and try to forsee what will happen to those trends in the future. Key rate forecasting is carried out in fairly <u>ad hoc</u> manner and several scenarios are mapped out. High, middle high, middle low and low scenarios are proposed and the projections conditional on those scenarios are worked out. The usual idea in formulating scenarios is to assume continued change in the short run along the same path as in the past followed by asymptotic shifting towards a long run equilibrium value. The long run values of the rates may be in the same direction as the observed short run trend or they may be in the opposite direction.

Formal time or space-time series analysis has been carried out on population stock numbers themselves (see, for example, Pearl and Reed, 1920, Pittinger, 1976 or Bennett, 1975a, 1975b). This works well as long as the underlying processes of change continue along smooth paths, but usually a components view of population change is taken (as outlined earlier in section 2) and the rates associated with the components are forecast separately.

Mortality rates have, in the last century, declined fairly smoothly with time in developed countries, and this mortality decline has spread outwards from North Western Europe and settled lands in North America and Australasia to other parts of the world. The decline was rapid as infectious diseases were conquered but has slowed considerably as the causes of death became concentrated on "endogeneous" diseases (heart disease, strokes, cancers) not clearly linked to an external infective agent. Negative exponential equations have been used to model recent mortality trends by age with some success (see O.P.C.S., 1978; Rees, 1977a for examples). Some modest improvement in mortality rates in Britain can still be expected because other countries have better mortality experiences. One mortality rate scenario is usually sufficient.

The same success has not attended fertility rate forecasting, and the major errors in British regional forecasts in the 1960's have been the result of getting the rates for this component wrong (Rees, 1980). The main problem has been a failure to predict when trends (upwards or downwards) will end and new ones occur. In the 1970's forecasters (O.P.C.S., 1972, 1973, 1974, 1975, 1976, 1977, 1978) have been at pains to predict an end to the post-1964 fall in fertility rates, although the turning point had to be postponed for several years. In 1978 this approach was vindicated (see O.P.C.S., 1979, Table 9 for the latest statistics).

Application of econometric time series techniques to the problem of fertility rate forecasting has not been very fruitful as Passel's 1976 extensive analysis demonstrates. The reason is that, since the end of the demographic transition in developed countries, populations have varied in their desire to have children in a fairly complex way, responding to long term shifts in social attitudes to the family and the role of women, to short term shifts due to economic conditions and to propaganda on behalf of family limitation (Simons, 1980). The forecaster is thus faced with the daunting task of predicting all these influential variables in order to forecast the fertility rate schedule.

When we turn to the forecasting of migration rates, a distressing situation reveals itself in Britain. For interregional migration rates the forecaster must depend on the Census as a source since the official continuous social survey - the General Household Survey - fails to give any regional breakdown of migration results. There are just three interregional migration tables of one year duration (1960-1, 1965-6, 1970-1) and two interregional migration tables of five years duration (1961-66, 1966-71) upon which to base conclusions as to trend. Comparisons are hampered by substantial and repeated changes in the areal units of measurement over the period.

Some consistent evidence on the rates of interregional migration is assembled in Table 9 from statistics given in Department of the Environment (1971) and Rees (1979a). The rate of interregional migration within Great Britain can be measured in 1960-61, 1965-66 and 1970-71 for the standard planning regions. The rate increases steadily through the decade from 12.3/1000 in 1960-61 to 13.3-13.5/1000 in 1965-66 to 15.4-15.5/1000 in 1970-71, an increase of 8% in the first quinquennium and 14% in the second quinquennium. A 14% increase was also noted by Stillwell (1979, Chapter 9) in the interregional migration rates from 1961-66 to 1 56-71. The apparent levelling of the trend in the 1970's may well be spurious since the 1975-76 figures are derived from accounts updated from 1970-71 by RAS type methods using row and column constraints from the 1975-76 period.

Table 9. Migration rates, Great Britain, 1960-61, 1965-66, 1970-71 ("old" regions); 1970-71, 1975-76 ("new" regions)

RATES

One year period	Regions	Internal out-migration or interregional migration rate	External out-migration or emigration rate	Total out-migration rate	External in-migration or immigration rate
1960-61	OR	.0123	-	-	
1965-66(a)	OR	.0133	-	-	
1965-66(b)	OR	.0135	(.0087)	.0222	.0064
1970-71(a)	OR	.0154	(.0088)	.0242	.0071
1970-71(b)	NR	.0155	(.0081)	.0236	.0070
1975-76	NR	(.0156)	.0068	(.0224)	.0064

Notes

1. The 1960-61 and 1965-66(a) rates are worked out from figures given in the Department of Environment (1971), Appendix 2 (interregional migrants) and Table 1.9 (mid-year populations 1960 and 1965). The immigrant numbers were derived from the Census 1961 and Sample Census 1966 migration tables.

2. The 1965-66(b), 1970-71(a), 1970-71(b) and 1975-76 rates are worked out from statistics given in Tables 7, 8, 9, 10 and 11 of Rees (1979a).

3. The internal and external in-migration migrant numbers derive from the 1966 Sample Census or 1971 Census migration tables, except for the internal 1975-76 figures which are estimated by applying RAS updating to the 1970-71 figures. The emigrant numbers and the immigrant numbers in 1975-76 are in all cases estimates based on adjustment of International Passenger Survey Statistics as described in Rees (1979b). The IPS statistics derived are for the period indicated.

4. Only exist-survive migrants are used as the rate numerators. The populations used as denominators are the start of period populations in all cases.

5. OR = "old" regions = regions current at time of the 1966 Sample Census and and 1971 Census.
NR = "new" regions = regions as defined subsequent to local government reorganization on 1 April 1974.

6. The brackets indicate less reliable estimates.

The picture is rather altered when external migration is considered. Here survey information is available for each year. The emigration rate remains fairly constant between 1965-66 and 1970-71 but drops 1975-76. The differences between the "old" region and "new" region figure indicates what can happen when two separate methods of estimation of this external flow are employed on different occasions. Immigration rate estimates are more reliable and these show a rise followed by a fall. The gap between emigration and immigration rates was in favour of emigration in all years but shows a narrowing from -2.3/1000 in 1965-66 to -0.4/1000 in 1975-76.

This rate series could be refined for external migration, but for internal migration between regions or local areas in the United Kingdom there exists no reliable series. The last reliable figures we have are those for 1970-71 and 1966-71 in the 1971 Census. This is an extremely unsatisfactory situation: no quantitative analysis is possible without "quantities". Further analysis of migratory trends will have to await the 1981 Census.

The evidence on the percentages of heads of household making one or more moves over the five years prior to the survey in Great Britain available from the General Household Survey (O.P.C.S., Social Survey Division, 1973, Table 5.52; 1975, Table 2.43; 1976, Table 2.24; 1977, Table 2.29 and 1978, Table 4.32) is of no comparable shift upwards in the migration rate. The percentage in 1971 was 35, in 1972, 36, in 1973 and 1974, 33 and in 1975, 32. The figures indicating a lessening general mobility.

The lack of a proper annual time series of internal migration figures is probably the key cause of lack of interest revealed in Woodhead (1979) in the adoption of multiregional models among British local planning authorities. This lack contrasts strongly with the availability of such series in a number of European countries - the Netherlands, Denmark, Sweden, Norway, Finland, the

Federal Republic of Germany, Czechoslovakia, Hungary and Bulgaria. These countries all have registration systems requiring residents to report changes of address. Such a system is probably a political non-starter in this country but at least some information could be provided from the General Household Survey of O.P.C.S. by further disaggregation of tables on household migration which would radically improve the RAS constrained estimates of accounts for inter-censal years.

Time series analysis can thus be said to have played a successful role in studying the mortality component, an unsuccessful role in fertility investigation but one with plenty of potential for further exploration (of the Easterlin and Condran (1976) hypothesis, for example). No role has yet been possible for the migration component in the U.K.

5.4 Can population projections be tested?

In formal time series analysis the time series model is always tested. This is rarely done with the population projection models discussed in sections 3 and 4. A cross-sectional test is often carried out of migration models involved (as in Masser, 1976, but not in Gilje and Campbell, 1973). A cross-sectional test of the multiregional cohort survival type model yields information only on the degree of rounding produced by the number of decimal places adopted in the rates, rather than a test. Stillwell (1979, Chapter 9) has carried out a number of migration model tests using 1961-66 based data to predict 1966-71 migration flows. One of the key elements in the projection that has to be got right is the overall change in the rate of migration in the system - no matter how sophisticated the distribution model it will not solve this problem. Good projections therefore depend on having a good series of migration data.

So, in this important aspect of applied population geography reliance must be placed on "after the fact" testing of projections - did the projections of 5 years ago turn out to be true? - rather than "before-the-fact" testing with a good historical series. Several authors have suggested that it would be useful to planners if forecasters assigned *a priori* probabilities to rate projections and thence to population projections. One might, for example, adopt a mean long term total fertility rate (TFR) of 2.1 and assume that the range 1.8 to 2.4 represented the range minus and plus one standard deviation about the mean. The probability that the long term TFR would fall below 1.8 would therefore be 0.16, for example. These probabilities might be based on a survey of informed opinion (the Delphi technique) or on a careful survey of family intentions.

The uncertainties associated with population projections will, of course, tend to vary with the age groups being considered and the future period of projection, as the Central Policy Review Staff (1976) have shown for the national projections. However, their conclusions may not carry one to sub-national areas because of variability in migration.

Critical investigation of the issue of testing population projections is indicated for the 1980's using the new migration data of the 1981 Census.

6. Why do people move around?

6.1 Introduction

So far the discussion of population change and movement has focussed on measurement and model representation issues. This emphasis has been adopted because the issues of measurement and formal model representation have attracted much less attention than they probably deserve. Before we can understand a phenomenon we must properly measure it.

However, researchers have always been eager to know the reasons for measured patterns and perhaps the most popular activity among population geographers in the 1970s and earlier has been the construction and cross-sectional testing of migration models. This field has been reviewed many times over the past decade from a variety of different viewpoints by Weeden (1973), Shaw (1975), Willis (1974), Stillwell (1975 and 1979, Chapter 6) and Gleave and Cordey-Hayes (1977). Detailed mathematical frameworks have been proposed by Alonso (1973, 1978), Wilson (1970, 1971, 1974, 1980) and investigated in detail by Ledent (1978). Remarks here will therefore be brief, and suggestive rather than definitive. To go further would be to reveal this author's ignorance.

Three frameworks appear to characterize the migration modelling field:

(i) the Markovian framework;
(ii) the spatial interaction framework; and
(iii) the econometric framework.

These can, of course, be linked. Alonso (1973, 1978) integrates the Markovian and spatial interaction approaches. Wilson (1980) shows that Alonso's mathematical framework is very close to the one developed earlier in the transport field (Wilson, 1970, 1971, 1974). In this framework econometric modelling of origin interaction propensities and destination attractiveness has been carried out.

6.2 The Markovian framework: the characteristics of movers

The population models described in sections 3 and 4 depend on the Markov assumptions that transitions between states in a population depend only on the current state (the single step assumption), and that the probabilities of transition apply uniformly to all the population in the current state (the homogeneity assumption). These assumptions are often untrue. In section 4 we saw that the obvious solution-expansion of the state-space (distinguishing smaller population groups in which the homogeneity assumption was felt to hold) - led to problems of 'dimensionality' that required solution by means of aggregation or decomposition. State-space expansion can in certain circumstances help relax the single step assumption but the discussion of this is postponed until life tables are considered.

Characteristics of the population that are held to affect the propensity to migrate include sex, age, marital status, educational attainment, occupation, income, and racial or ethnic status.

Sex differences in migration propensity are systematic but largely explained by the differing age compositions of male and female population. The male and female migration rates are very similar except that the female schedule precedes the male in the labour force ages by the average age gap between spouses (Plessis-Fraissard, 1977).

The schedule of migration rates by age reflects the association of migrations with significant transitions in the life cycle. In childhood the migration rates are linked to and dependent on the adult migration rates. One observation that has been made is that the parental shift (the age gap between the childhood and adult curves at the same migration rate level) widens as the child gets older. This is not to be explained in the Great Britain case (see Figure 6) by changes in the average age of maternity/paternity (which moved in a direction opposite to the change in parental shift observed). It is probably due to a divergence in the migration behaviour of adults with children and those without - a hypothesis for further investigation.

As the child becomes an adult in his/her late teens, migrations occur as first jobs or careers are taken up, or as further education is begun and finished, as marriage partners are found, or if the armed forces are entered. These life cycle events will be crowded in the 17-24 years, and will become less frequent with age. The raising of a family will tend to make a parent less locationally mobile as there are more activity spaces to be shifted on each migration. The family will also tend to have moved into the more secure parts of the housing market (owner-occupation or council tenure) where the pressures to move are less and the costs higher. Mobility will rise again briefly in the ages around retirement, particularly in streams to favoured amenity regions.

The different socioeconomic groups will also show marked differences in their mobility propensities. Kiernan (1980) and Holcraft (1980) have studied the characteristics of persons moving interregionally in a longitudinal cohort study and emphasise the radically higher mobility experiences of those with A-levels or further education qualifications. This mobility is associated both with the need to move to gain that education and also with the further flung job market for the specialist skills acquired in further education. Occupational differences in migration propensity are highlighted by Gleave and Cordey-Hayes (1977) and Gleave and Palmer (1980), with a clear association of interregional migration rates with occupational ranking and with employee (more migratory)/ employer (less migratory) status. Given the differences between education attainment and occupational groups in migration behaviour, similar differences should characterize the different income groups, were we able to observe such interregional migration rates although the evidence is that the differentials are much weaker (OPCS, Social Survey Division, 1973, Tables 5.53 and 5.54; Bogue, 1969, Chapter 19, Table 19-12).

Can we use these population characteristics to improve our Markov-type models of migration? We can, as long as we are prepared to pool information and not use all the inter-state transitions that result. A simple example will illustrate some of the possibilities.

Assume that the population stock in a region, classified by age, sex and marital status is available, P_r^{ixm}, where m stands for marital status (single, married, widowed, divorced) then the number of out-migrants M_r^i might be modelled as

$$M_r^i = \sum_{m,x} P_r^{ixm} \; m_r^{ixm\,\mathrm{OUT}} \qquad (92)$$

and the interregional migration as

$$M^{ij}_{r\;r+1} = M_r^i \; \hat{p}\,(j, r+1/i, r) \qquad (93)$$

and for $\hat{p}$ (j, r + 1/i, r), the modelled probability of surviving the period in region j in age group r + 1 given an initial location in region i and an initial age group, instead of using the observed probability from a set of accounts a term could be borrowed from the spatial interaction framework (Gleave and Cordey-Hayes, 1977, p 47)

$$\hat{p}\,(j, r+1/i, r) = \hat{p}\,(j/i) = \frac{P_j q_j f(c_{ij})}{\sum_{k=1}^{n} P_k q_k \, f(c_{ik})} \qquad (94)$$

where P_j is the population of the destination region, q_j an index summarizing the attractiveness of region k to migrants and $f(c_{ij})$ some function of the separation of region i from region j.

The model used by Gleave and Cordey-Hayes (1977) differs from the simple one outlined above in using 'duration of residence' as the key population characteristic determining propensity to move. Unfortunately in so doing they employ data in which the correct probabilities are not identified and use arguments in which empirical, verifiable hypotheses are raised to the status of axioms. The problem of incorrect definition of migration rates or probabilities by duration of residence category has plagued this subfield and it is with the work of Courgeau (1973), Plessis-Fraissard (1977, 1979) Clark and Huff (1977), Hobcraft (1980) and others that measurement problem has been clarified.

The problem arises from a confusion of population-at-risk with prospective probability. Figure 11 illustrates a typical situation. The number of life-lines in each duration of residence category at time t is as follows:-

Duration of residence category		Population count	Distribution proportion	Lifelines
d_1	less than one year	$P_{d_1} = 8$	.57	7,8,9,10,11,12, 13,14
d_2	less than two years but more than one year	$P_{d_2} = 4$	.29	3,4,5,6
d_3	less than three years but more than two years	$P_{d_3} = 2$	.14	1,2
	Total	14	1.00	

However, the probabilities of making a transition between time t and t + 1 are as follows:-

	Population at risk P_{d_i}	Number of transitions T_{d_i}	Probability $p\ (T/d_i)$
d_1	8	4	.5
d_2	4	2	.5
d_3	2	1	.5

The distribution proportions of the population at risk across the duration categories have often been confused with the probability of making a transition given that a person was observed to be in a duration category. This type of prospective probability is the one required in the Gleave and Cordey-Hayes model.

Courgeau (1973) and Plessis-Fraissard (1977, 1979) made careful analyses of samples of migration histories and came to the conclusion that, if age was controlled, the probabilities of migration were independent of duration of residence: that is, that

$$p\ (T/d_1,\ r) \ = \ p\ (T/d_2, r) \ = \ \ldots\ p\ (T/d_n,\ r) \qquad (95)$$

for all age groups r, with the possible exception of the ages immediately

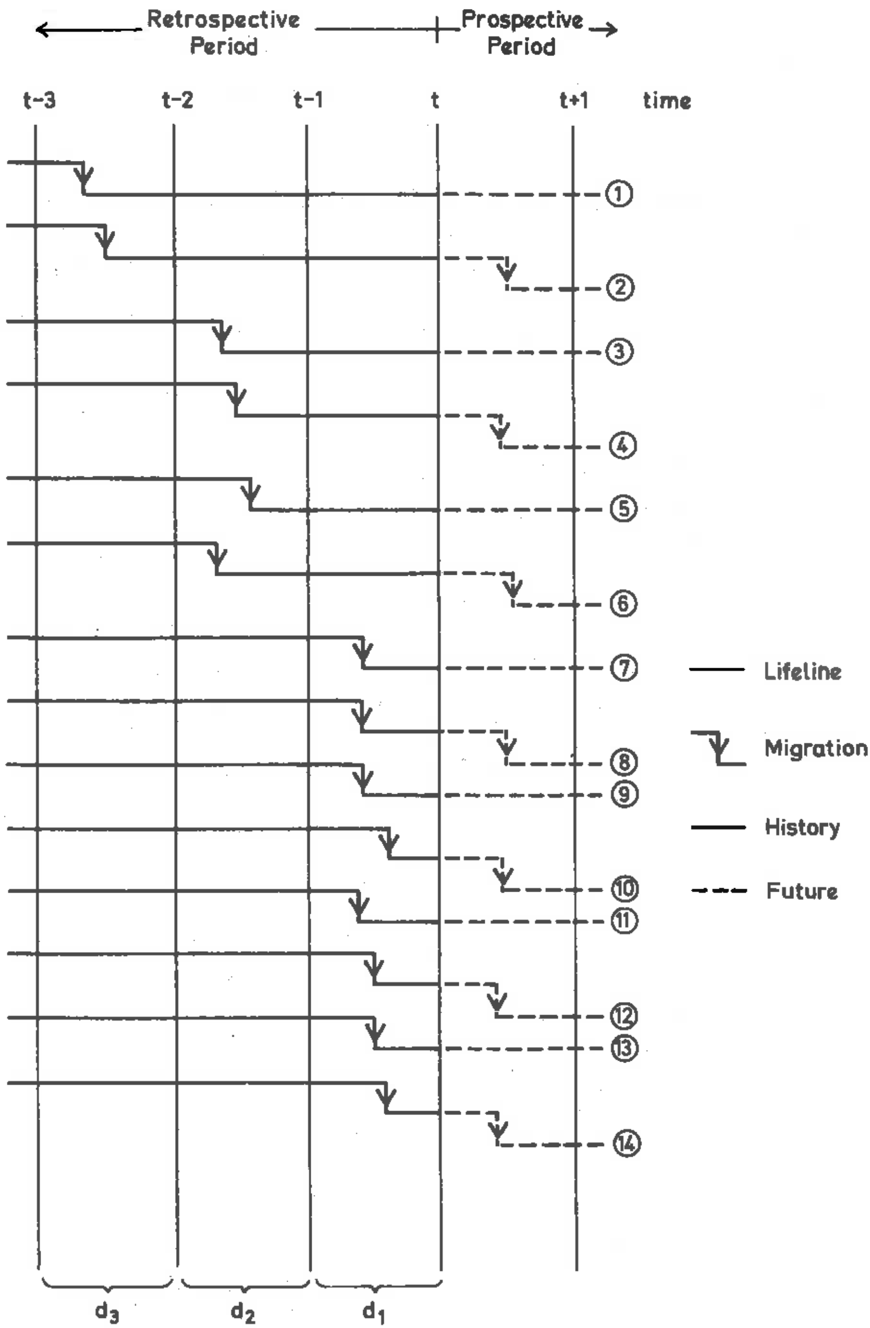

Figure 11. A sample of lifelines illustrating the problem of measuring the probability of migration conditional on duration of residence

around age 20. Even in that situation the dependence was not marked.

6.3 The spatial interaction framework: deviations from the transport model experience

Gravity and intervening opportunity models have long been used in the migration modelling field but it is in the context modelling of the daily journey to work that a robust mathematical framework has been developed (Wilson, 1974). Instead of describing the many studies that implement members of the spatial interaction model family (see the reviews cited earlier for this), an attempt is made here to compare models in the two fields in order to gain insight into the ways in which migration models might be improved.

The typical variable to be predicted in a disaggregated transport model might be

T_{ij}^{kpn} = trips of purpose p by person type n on mode k between origin zone i and destination zone j

The equivalent migration variable might be

$M_{r\ r+1}^{ijx}$ = migrants of age transition r to r + 1, sex x between origin zone i and destination zone j.

The spatial labels are equivalent though the migration zones are likely to be much further apart, of course; the age group and sex categories correspond with the person type. Mode in the transport model has no relevance in the migration context. But what about the index p in the transport model, which refers to journey to work or shopping or school or recreation or business trip purposes? The different trip purposes are modelled separately.

The equivalent of purpose in the migration model should be 'reason' or 'cause' of migration. However, in virtually all countries bar Czechoslovakia (Rogers and Castro, 1979) no comprehensive migration statistics classified by purpose or reason are available. This creates difficult problems in model building since the propensity and attractiveness indexes will differ radically from one migration purpose to another. For example, job-oriented migrants will assess potential destinations in terms of the job opportunities they offer; retirement migrants will be influenced by bungalow availability, amenity provision and climate. Persons moving for reasons relating to housing will be concerned with the characteristics and costs of housing at various locations.

Various attempts have been made to use surrogate classifications to achieve this key classification by reason for migration. Attention is restricted to migration in the age range 20-64 in labour force oriented models (Gleave and

Cordey-Hayes, 1977, for example). In other studies, eg.Masser, 1970) only migration flows among widely separated areas (conurbations) are considered. Intra-region or intra-zone migration is routinely excluded from interregional migration models. A migration between contiguous areas and non-contiguous areas is modelled separately (Stillwell, 1979) or by means of a contiguity dummy (Weeden, 1973). Attempts (Stillwell, 1979) are made to split aggregate migration matrices into several purpose components by application of probabilities of migration by reason by distance land (Harris and Clausen, 1966; Weinstein, 1975) or by elaborate split model calibration exercises (Hyman and Gleave, 1978). All these studies contribute to our knowledge of migration, but all authors would, I feel sure, trade in their elaborate models for British equivalents to the Czech data (detailed interdistrict migration tables classified by reason for moving for each year). If reason-for-moving data were available it would be much easier to specify precise variables for the mass terms in gravity models or the attractiveness factors in spatial interaction models.

When the impedance variable in migration is examined it is clear that migration modellers have come up with no equivalent of the transport modellers' generalized cost equations. Usually airline or road distance is employed in a fairly mechanical way. Distance in this case is probably not acting as a surrogate for transfer cost since this probably pays a much less important part in migration than in intra-urban daily movement and since this has a much larger fixed cost element. Hägerstrand (1957) has argued that it is the decay of information about opportunities with distance that is important, and that this decay need not be exclusively associated with distance. Levels of explanation in migration models might be improved by substituting an information flow measure (telephone call traffic or postal traffic) for distance. This would itself need modelling, of course.

The distance measure in migration models does not have the same policy significance as the cost function in transport models. In projecting the migration pattern into the future (a rare exercise but see Stillwell, 1979 and Gleave and Cordey-Hayes, 1977) using a spatial interaction model, the distance impedance matrix either remains unchanged or changes marginally. Hence use of a simpler constrained growth factor model (resembling an RAS type updating model), which gives a better fit than more complex models containing distance functions (Stillwell, 1979) may still be appropriate.

On the other hand such a conclusion would tend to discourage the experimentation within a proper spatial model framework of the a distance functions proposed by migration modellers (Morrill, 1965; Taylor, 1971). There remains also the problem of projecting the origin and destination in- and out-migration totals, which we saw in Table were on a steady increase in Britain, following perhaps Zelinsky's mobility transition. predictions (Zelinsky, 1971).

6.4 A comparison of Markovian and spatial interaction models

Detailed comparison of the long run behaviour of the linear (Markovian) population model and non-linear models incorporating spatial interaction model features has been carried out by Ledent (1978c). He carries out a detailed mathematical investigation of the existence and convergence properties of a long run equilibrium of the non-linear model. The conclusion that such several equilibrium states of a multiregional system may exist, that they may be dependent on the initial system state, and that one or more zero states may exist at equilibrium. These features he argues make the non-linear model unsatisfactory compared with the linear model for stable state investigations.

6.5 The econometric framework and demoeconomics

Such a conclusion would, of course, be unacceptable to those attempting to verify hypotheses about migration behaviour derived from economic theory. A continual search for better regression models characterizes much of the 'regional science' effort in migration study. Much of this effort might benefit from embedding the econometric equations describing the 'attractiveness' of destinations or 'repulsiveness' of origins in a consistent spatial interaction framework.

However, the importance of making robust connections between the population system and the economy is such that fruitful interaction of this style of modelling and that of the two other frameworks is bound to develop. A vigorous literature in the interface between demographic and economic studies is emerging, and is extensively and perceptively reviewed in Ledent (1978d) with recent British examples with an applied planning orientation being Madden (1976) and Breheny and Roberts (1978).

7. How long do people live and where?

Before closing this already lengthy review about one of the major achievements of the past we can now begin go add to traditional answers people live?' some notion of where their lives

The long established life table model of d expanded into a multiregional life table model developed as an operational tool by Ledent (197 definitions) by Willekens and Rogers (1976, 197 by Rees and Wilson (1975, 1977) and by Rees (19 and Willekens (1979) (application to working li have been applied in a variety of countries by tion with the International Institute for Appli Tekse, 1978; Drewe, 1978; Koch and Gatzweiler 1979a; Rikkinen, 1979 and Termote, 1978).

Application of the methods has revealed - some of the difficulties of the underlying Mark include dependence of migration probabilities o because of return and repeat migration (Rees, 1 1977); the dependence of current migration prol (Long and Hansen, 1975), possibly to be solved l and the problem of converting movement data int (Ledent, 1979), perhaps insoluble. These probl faced in population projection methodology exce form in multistate life tables, for the simple are expressed in matrix form, for example, the place of residence and by place of birth, rathe population projections, populations by place of

8. Conclusion

What then, have been the main achievements of the past decade in population geography and what are the problems it still faces?

The theme that runs through work on population geography is a concern to account for, to build into models and to explain the interactions (migration) among regions within nations. A common belief was explored that these explicit connections had to be represented in the system models that were built. Success in this aspect of population studies probably meant that less attention was paid to the interactions between the population system and other systems involving human activity, especially the economy and the housing market. The neglect of inter-system interactions compared with interregion interactions could, however, be regarded simply as an 'apparent' effect, a product of the neglect of the reviewer. To repair this neglect or this defect is one of the challenges for population geography in the 1980s.

As the reader was warned in the introduction this review has strayed far beyond the traditional concerns of population geography, although much more could have been said on all the topics covered here. And yet the list of research areas cursorily covered in or omitted from this review - marriage patterns, simulation of populations, household patterns and formation processes, population policies ... - is legion, and testimony surely to the strength of interest in population matters. Problems in model construction and testing, data gathering and empirical analysis abound, but the variety of contribution from sundry disciplines surely represents a good omen for the coming decade. Population geographers should feel privileged to be part of this ferment.

References

Alonso, W. (1973) National interregional demographic accounts: a prototype. Monograph No.17, Institute of Urban and Regional Development, University of California, Berkeley.

Alonso, W. (1978) A theory of movement. In Hansen, N.M. (ed.) Human Settlement. Ballinger, Cambridge, Mass.

Bacharach, M. (1970) Biproportional matrices and input-output change. Cambridge University Press, Cambridge.

Baxter, R. and Williams, I. (1978) Population forecasting and uncertainty at the national and regional scale. Progress in Planning, 9, 1, Pergamon Press, Oxford.

Bennett, R.J. (1975a) Dynamic systems modelling of the North West region: 1. spatio-temporal representation and identification. Environment and Planning A, 7, 525-538.

Bennett, R.J. (1975b) Dynamic systems modelling of the North West region: 2. estimation of the spatio-temporal policy model. Environment and Planning A, 7, 539-566.

Bies, K. and Tekse, K. (1978) Migration and settlement in Hungary. Working Paper WP-78-20, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Bogue, D.J. (1969) The principles of demography. John Wiley, New York.

Brass, W. (1974) Perspectives in population prediction: illustrated by statistics of England and Wales. Journal of the Royal Statistical Society, A, 137, 4, pp.532-570.

Brown, L.R. (1976) World population trends: signs of hope, signs of stress. World Watch Paper 8, World Watch Institute, Washington, D.C.

Breheny, M.J. and Roberts, A.J. (1978) An integrated forecasting system for structure planning. Town Planning Review, 49, 3, 306-318.

Byron, R. (1978) The estimation of large social account matrices. Journal of the Royal Statistical Society, A, 141, 3, 359-367.

Campbell, R.A. (1976) Local population projections. Population Trends, 5, 9-12.

Central Policy Review Staff (1976) Population and the social services. H.M.S.O., London.

Champion, A.G. (1976) Evolving patterns of population distribution in England and Wales. Transactions, Institute of British Geographers, New Series, 1, 401-420.

Clark, W.A.V. and Huff, J.O. (1977) Some empirical tests of duration-of-stay effects in intra-urban migration. Environment and Planning A, 9, 1357-1374.

Compton, P.A. (1969) Internal migration and population change in Hungary between 1959 and 1965. Transactions, Institute of British Geographers, 47, 111-130.

Compton, P.A. (1978) Fertility differentials and their impact on population distribution and composition in Northern Ireland. Environment and Planning A, 10, 12, 1343-1456.

Cordey-Hayes, M. (1975) Migration and the dynamics of multiregional population systems. Environment and Planning A, 7, 793-814.

Courgeau, D. (1973) Migrants et migrations. Population, 28, 1, 95-129.

Department of the Environment (1971) Long term population distribution in Great Britain - a study. H.M.S.O., London.

Department of the Environment (1976) British cities: urban population and employment trends 1951-71. Research Report 10. Department of the Environment, London.

Drewe, P. (1978) Migration and settlement in the Netherlands: dynamics and policy. Working Paper WP-78-17, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Easterlin, R. and Condran, G. (1976) A note on the recent fertility swing in Australia, Canada, England and Wales, and the United States . In Richards, H. (ed.) Population, factor movements and economic development: studies presented to Brinley Thomas, University of Wales Press, Cardiff.

Eversley, D.E.C. (1971) Population change and regional policies since the war. Regional Studies, 5, 211-228.

Gilje, E.K. and Campbell, R.A. (1973) A new model for projecting the population of the Greater London boroughs. Research Memorandum 408, Greater London Council, London.

Girling, P.D. (1979) Developing the migration component of the official sub-national projections: specification of research provided with invitation to tender. Unpublished document, Department of the Environment, London.

Gleave, D. and Cordey-Hayes, M. (1977) Migration dynamics and labour market turnover. Progress in Planning, 8, 1, 1-95. Pergamon Press, Oxford.

Gleave, D. and Palmer, D. (1980) The relationship between geographic and occupational mobility in the context of regional economic growth. In Hobcraft, J.N. and Rees, P.H. (ed.) Regional demographic development, Croom Helm, London.

Hägerstrand, T. (1957) Migration and area. In Hannerberg, D. (ed.) Migration in Sweden. Lund Studies in Geography, Series B, No.13.

Harris, A.I. and Clausen, R. (1966) Labour mobility in Great Britain, 1953-63. H.M.S.O., London.

Hobcraft, J.N. (1980) Forecasting the components of population change, with special reference to migration. In Hobcraft, J.N. and P.H. Rees (ed.) Regional demographic development. Croom helm, London.

Hyman, G. and Gleave, D. (1978) A reasonable model of migration. Transactions, Institute of British Geographers, New Series, 3, 2, 179-201.

Illingworth, D.R. (1976) Testing some new concepts and estimation methods in population accounting and related fields. Ph.D. dissertation, School of Geography, University of Leeds.

Jackson, D. (1979) Current population research in the United Kingdom: a classified inventory of current research projects and selected bibliographies on the consequences of contemporary population trends. CPS Working Paper No.79-2, Centre for Population Studies, London School of Hygiene and Tropical Medicine.

Jenkins, J.C. and Rees, P.H. (1977) Computer programs and aggregate accounts-based and associated forecasting models of population. Working Paper 205, School of Geography, University of Leeds.

Jones, H. (1975) Spatial analysis of human fertility in Scotland. Scottish Geographical Magazine, 91, 102-113.

Kennett, S. (1977) Migration trends and their contribution to population trends in the urban system (1961-1971). Working Report No.50, Urban Change Project, Department of Geography, London School of Economics.

Kennett, S. (1980) Migration within and between the metropolitan areas of Britain, 1966-71. In Hobcraft, J.N. and Rees, P.H. (ed.) Regional demographic development. Croom Helm, London.

Keyfitz, N. (1977) Applied mathematical demography. John Wiley, New York.

Keyfitz, N. (1979) Comment: multidimensional demography and its data. Environment and Planning A, 12, forthcoming.

Kiernan, K. (1980) Characteristics of young people who move interregionally: a longitudinal study. In Hobcaft, J.N. and Rees, P.H. (ed.) Regional demographic development. Croom Helm, London.

Koch, R. and Gatzweiler, H-P. (1978) Migration and settlement in the Federal Republic of Germany. Working Paper WP-78-39, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Lawton, R. (1977) People and work. In House, J.W. (ed.) The U.K. space: resources, environment and the future. Second edition. Weidenfeld and Nicholson, London.

Lawton, R. (1980) Regional population trends in England and Wales, 1750-1971. In Hobcraft, J.N. and Rees, P.H. (ed.) Regional demographic development. Croom Helm, London.

Ledent, J. (1978a) Some methodological and empirical considerations in the construction of increment-decrement life tables. Research Memorandum RM-78-25, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Ledent, J. (1978b) Temporal and spatial aspects in the conception, estimation and use of migration rates. Paper presented at the IIASA Conference on Analysis of Multiregional Population Systems: Techniques and Applications, September 19-22, 1978.

Ledent, J. (1978c) Stable growth in the non-linear components-of-change model of interregional population growth and distribution. Research Memorandum, RM-78-28, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Ledent, J. (1978d) Regional demoeconomic growth: a survey of theories and empirical models. Working Paper, WP-78-51, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Ledent, J. (1979) Multistate life tables: movement versus transition perspectives. Environment and Planning A, 12, forthcoming.

Liaw, K-L. (1977) Discrete-time sex- and age-disaggregated spatial population model: theoretical analysis and Canadian case study. Research report to the Canada Council, Department of Geography, McMaster University, Hamilton, Ontario.

Liaw, K-L. (1978a) Dynamic properties of the 1966-1971 Canadian spatial population system. Environment and Planning A, 10, 389-398.

Liaw, K-L. (1978b) Sensitivity analysis of discrete-time, age-disaggregated interregional population systems. Journal of Regional Science, 18, 263-281.

Liaw, K-L. (1979) Multistate dynamics: the convergence of an age-by-region population system with a time invariant structural matrix. A Canadian case study. Environment and Planning A, 12, forthcoming.

Long, L.H. and Hansen, K.A. (1975) Trends in return migration to the South. Demography, 12, 4, 601-614.

Long, L.H. and Hansen, K.A. (1977) Interdivisional primary, return, and repeat migration. Public Data Use, 5, 2 , 3-10.

Macgill, S.M. (1975) Balancing factor methods in urban and regional analysis. Working Paper 124, School of Geography, University of Leeds.

Macgill, S.M. (1977) Activity-commodity spatial interaction models and related applications. Ph.D. thesis, School of Geography, University of Leeds.

Madden, M. (1976) A multilevel forecasting framework. Unpublished paper, Department of Civic Design, University of Liverpool.

Masser, I. (1970) A test of some models for perdicting inter-metropolitan movement of population in England and Wales. Working Paper No.9, Centre for Environmental Studies, London.

Masser, I. (1976) The design of spatial systems for internal migration analysis. Regional Studies, 10, 39-52.

Masser, I. and Brown, P.J.B. (1975) Hierarchical aggregation procedures for interaction data. Environment and Planning A, 7, 509-523.

Mohs, G. (1979) Migration and settlement in the German Democratic republic. Working Paper WP-79-17, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Morrill, R.L. (1965) Migration and the spread and growth of urban settlement. Lund Studies in Geography, Series B, Human Geography, No.26.

O.P.C.S., Social Survey Division (1973) The general household survey: an introductory report. H.M.S.O., London.

O.P.C.S., Social Survey Division (1975) The general houshold survey, 1972. H.M.S.O., London.

O.P.C.S., Social Survey Division (1976) The general household survey, 1973. H.M.S.O., London.

O.P.C.S., Social Survey Division (1977) The general household survey, 1974. H.M.S.O., London.

O.P.C.S., Social Survey Division (1978) The general household survey, 1975. H.M.S.O., London.

O.P.C.S. (1972) Population projections No.2 1971-2011. H.M.S.O., London.

O.P.C.S. (1973) Population projections No.3 1972-2012. H.M.S.O., London.

O.P.C.S. (1974) Population projections No.4 1973-2013. H.M.S.O., London.

O.P.C.S. (1976) Population projections, 1974-2014. Series PP2, No.6. H.M.S.O., London.

O.P.C.S. (1977) Population projections, 1975-2015. Series PP2, No.7. H.M.S.O., London.

O.P.C.S. (1978) Population projections, 1976-2016. Series PP2, No.8. H.M.S.O., London.

O.P.C.S. (1979) Population trends 17, Autumn 1979. H.M.S.O., London.

Passel, J. (1976) Population projections utilizing age-specific and age-parity -specific birth rates predicted with time series models: projections with confidence limits. Ph.D. dissertation, The Johns Hopkins University. Thesis No. 76-22,942, Xerox University Microfilms, Ann Arbor, Michigan, 48106.

Pearl, R. and Reed, L.J. (1920) On the rate of growth of the population of the United States since 1790 and its mathematical representation. Proceedings of the National Academy of Sciences, 9.

Philipov, D. (1978) Migration and settlement in Bulgaria. Environment and Planning A, 10, 593-617.

Pittenger, D.B. (1976) Projecting state and local populations. Ballinger, Cambridge, Mass.

Plessis-Fraissard, M. (1977) Age, durée de séjour et probabilité de migrer. In L'analyse démographique et ses applications. V^e Colloque National de Démographie du Centre National de la Recherche Scientifique, No.934. Éditions du C.N.R.S., Paris. Pp. 193-211. (English version: Working Paper 107, School of Geography, University of Leeds.)

Plessis-Fraissard, M. (1979) Migrants, spouses and households: an investigation. Ph.D. dissertation in preparation, School of Geography, University of Leeds.

Plessis-Fraissard, M. and Rees, P.H. (1976) A computer program for constructing age-sex disaggregated multiregional population accounts (DAME): a full description. Working Paper 169, School of Geography, University of Leeds.

Rees, P.H. (1976) Modelling the regional system: the population component. Working Paper 148, School of Geography, University of Leeds. Revised version published in Clarke, J.I. and Pelletier, J. (ed.) (1978) Régions géographiques et régions d'aménagements. Éditions L'Hermès, Lyon.

Rees, P.H. (1977a) The future population of East Anglia and its constituent counties (Cambridge, Norfolk and Suffolk). A report prepared for the East Anglia Economic Planning Council by the University of Leeds Industrial Services Ltd. under Department of the Environment Contract Number DGR/461 /23.

Rees, P.H. (1977b) The measurement of migration from census data and other sources. Environment and Planning A, 9, 247-272.

Rees, P.H. (1978a) Problems of multiregional population analysis: data collection and demographic accounting. Working Paper 229, School of Geography, University of Leeds. Forthcoming in Sistemi Urbani.

Rees, P.H. (1978b) Increment-decrement life tables: some further comments from a demographic accounting point of view. Environment and Planning A, 10, 705-726.

Rees, P.H. (1979a) Migration and settlement: 1. United Kingdom. Research Report RR-79-3, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Rees, P.H. (1979b) Multistate demographic accounts: measurement and estimation procedures. Environment and Planning A, 12, forthcoming.

Rees, P.H. (1979c) Regional population projection models and accounting methods. The Journal of the Royal Statistical Society, Series A, 142, 2, 223-255.

Rees, P.H. (1980) Population forecasts for British regions: a comparison. In Hobcraft, J.N. and Rees, P.H.(ed.) Regional demographic development. Croom Helm, London.

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

Rees, P.H. and Wilson, A.G. (1975) Accounts and models for spatial demographic analysis 3: rates and life tables. Environment and Planning A, 7, 199-231.

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

Rikkinen, K. (1979) Migration and settlement: 2. Finland. Research Report RR-79-9, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Rogers, A. (1966) Matrix methods of population analysis. Journal of the American Institute of Planners, 32, 40-44.

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

Rogers, A. (1969) On perfect aggregation in the matrix cohort-survival model of inter-regional population growth. Journal of Regional Science, 9, 417-424.

Rogers, A. (1971) Matrix methods in urban and regional analysis. Holden-Day, San Francisco.

Rogers, A. (1973) The mathematics of multiregional demographic growth. Environment and Planning, 5, 1, 3-29.

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

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

Rogers, A. (1976b) The comparative migration and settlement study: a summary of workshop proceedings and conclusions. Research Memorandum RM-76-01, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Rogers, A. and Castro, L. (1979) Migration age patterns: II. cause-specific profiles. Working Paper, WP-79-65, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Rogers, A. and Philipov, D. (1979) Multiregional methods for subnational population projections. Working Paper WP-79-40, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Rogers, A., Raquillet, R. and Castro, L. (1978) Model migration schedules and their applications. Environment and Planning A, 10, 475-502.

Shaw, R.P. (1975) Migration theory and fact: a review and bibliography of current literature. Bibilography Series Number Five. Regional Science Research Institute, Philadelphia.

Simons, J. (1980) Developments in the interpretation of recent fertility trends in England and Wales. In Hobcraft, J.N. and Rees, P.H. (ed.) Regional demographic development. Croom Helm, London.

Stillwell, J.C.H. (1975) Models of inter-county migration: a review . Working Paper 100, School of Geography, University of Leeds.

Stillwell, J.C.H. (1979) Migration in England and Wales: a study of inter-area patterns, accounts, models and projections. Ph.D. dissertation, School of Geography, University of Leeds, Leeds.

Stone, R. (1963) (ed.) Input-output relations, 1954-1966. Chapman and Hall, London.

Stone, R. (1965) A model of the educational system. Minerva, 3, 172-186.

Stone, R. (1971) Demographic accounting and model-building. O.E.C.D., Paris.

Stone, R. (1972) A Markovian education model and other examples linking social behaviour to the economy. Journal of the Royal Statistical Society, A, 135, 4, 511-543.

Stone, R. (1975a) Towards a system of social and demographic statistics. Department of Economic and Social Affairs, Statistical Office, Studies in Methods, Series F, No.18. United Nations, New York.

Stone, R. (1975b) Direct and indirect constraints in the adjustment of observations. Paper, Department of Applied Economics, University of Cambridge.

Stone, R., Champernowne, D.G. and Meadem J.E. (1942) The precision of national income estimates. The Review of Economic Studies, 9,111-125.

Taylor, P. (1971) Distance transformations and distance decay functions. Geographical Analysis, 3, 221-238.

Termote, M. (1978) Migration and settlement in Cana da: dynamics and policy. Working Paper WP-78-37, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Weeden, R. (1973) Interregional migration models and their application to Great Britain. In National Institute for Economic and Social Research (ed.) Regional papers II. Cambridge University Press, London.

Weinstein, E.T.A. (1975) The movement of owner occupied households between regions. Regional Studies, 9, 137-145.

Willekens, F. (1977) The recovery of detailed migration patterns from aggregate data: an entropy maximzing approach. Research Memorandum RM-77-58, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Willekens, F. (1978) The comparative migration and settlement study: background paper for the 1978 September conference. Working Paper WP-78-10, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Willekens, F. (1979) Multistate analysis: tables of working life. Environment and Planning A, 12, forthcoming.

Willekens, F., Por, A., and Raquillet, R. (1979) Entropy, multiproportional and quadratic techniques for inferring detailed migration patterns from aggregate data. Mathematical theories, algorithms, and computer programs. Working Paper WP-79-88, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Willekens, F. and Rogers, A. (1976) Computer programs for spatial demographic analysis. Research Memorandum RM-76-58, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Willekens, F. and Rogers, A. (1977) More computer programs for demographic analysis. Research Memorandum RM-77-30, International Institute for Applied Systems Analysis, Laxenburg, Austria.

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

Willis, K.G. (1974) Problems in migration analysis. Saxon House, Farnborough.

Wilson, A.G. (1970) Entropy in urban and regional modelling. Pion, London.

Wilson, A.G. (1971) A family of spatial interaction models. Environment and Planning, 3, 1-32.

Wilson, A.G. (1974) Urban and regional models for geography and planning. John Wiley, London.

Wilson, A.G. (1980) Comments on Alonso's "theory of movement". Environment and Planning A, 12, forthcoming.

Wilson, A.G. and Rees, P.H. (1974) Accounts and models for spatial demographic analysis 2: age-sex disaggregated populations. Environment and Planning A, 6, 101-116.

Woodhead, K. (1979) Computer programs for population projections: a study of current local authority practice and requirements. Paper, Department of Planning, Cambridgeshire County Council, Cambridge.

Zelinsky, W. (1971) The hypothesis of the mobility transition. Geographical Analysis, 61, 219-249.