Working Paper 27

A REVISED NOTATION FOR SPATIAL DEMOGRAPHIC ACCOUNTS AND MODELS

P. H. Rees

Paper prepared for a Special Issue of

Environment and Planning on Spatial

Demography

Department of Geography

University of Leeds

November, 1972

1. Why notation is important

In urban and regional analysis we have an inadequate conceptual grasp of many of the dynamic processes which we know are important in changing the main features of the urban and regional system. When we look at what we know ab ut the way population changes over time we find that this is especially true. A great deal of work has been devoted to the analysis of population but most of the best (Keyfitz, 1968) has been in a single region framework. Yet we know that a single region framework is inadequate and does not represent what goes on in the real world. The papers collected together in this issue of Environment and Planning report on work in progress that is aimed at developing demographic analysis in a spatial or multi-regional framework.

Mathematical definition of verbally or intuitively understood concepts and the exploration of the implications of these definitions is the approach adopted by the authors. Use of mathematics in this way helps to clarify problems, to solve them and to discover things that were not intuitively obvious to begin with.

One of the key steps in such a mathematical analysis of a set of concepts is to define a notation for the variables involved. Half the battle in many theoretical research endeavours is won if a powerful notation is defined at the outset. Once such a notation has been devised many important relationships fall into place very quickly.

At the beginning of work on a particular problem or set of problems there are almost as many notations as there are authors. After a while, however, the advantages of particular notations become evident, and these are then generally accepted in the field. In the field of spatial interaction modelling, for example, the notation developed in transportation planning work and used to great effect by Wilson (1970), Cripps & Foot (1969), Batty (1970) and others, has been generally adopted. This notation could also be used in a retrospective sense as well to analyze what papers previous to the notation were saying.

In spatial demographic analysis such a convergence in notation has not yet occurred. Even individual authors have used different notations as their work has developed. Compare the notation used by Rogers in his 1968 book to develop a multi-regional cohort survival model with that he presents in his paper in this issue of Environment and Planning. Compare the notations used in the Wilson (1972a) and Rees (1972) papers with those used in their joint paper in this issue. A period of rapid evolution of notation is occurring, and the juxtaposition in this issue of papers from a variety of authors on the same kinds of problems should aid greatly in this process.

What the juxtaposition has already suggested to the present author is that there is another and possibly more powerful notation for spatial demographic analysis which could be used. This notation borrows features from previous work by Rogers (1968), by Keyfitz (1968), by Wilson (1972a and 1972b) and by Rees (1972) and from the two papers by Rogers and by Rees and Wilson in this issue. This notation is described in detail below with the hope that it will contribute to notational convergence.

2. The basic variables and subscripts

In previous papers different events which happen to people (birth, death, migration, marriage) were sometimes represented by different variable letters, sometimes by different subscript lists. Let us here assign letters to variables that have the same meaning in terms of the actors involved and let us use subscripts to represent the different things that can happen to the actors.

Let

K represent population numbers

V represent time in length or a period of time, T being a particular value of this i, j   subscripts represent <u>locations</u>

T, S   subscripts represent <u>age</u> groups

X   represent <u>sex</u> (X = M is male, X = F is female)

t, t+T subscripts represent <u>points</u> in time

Most of the other variables that have to be defined in spatial demographic analysis are combinations of the variables defined here.

Let us assign the subscripts locations with respect to the population variable. A population flow through a variety of states could be represented as

$$\cancel{K^{i_1,\, i_2,\, i_3,\, \ldots,\, i_n X}_{r_1,\, r_2,\, r_3,\, \ldots,\, r_n} (t_1,\, t_2,\, t_3,\, \ldots,\, t_n)} \qquad K^{i=1,2,3,\ldots,N,X}_{r=1,2,3,\ldots,R}(t=1,2,3,\ldots,\lambda)$$

The locational subscripts are assigned the upper right position and are followed by the sex subscript. The age subscripts are assigned the bottom right position and time subscripts are listed in brackets following the variable. Subscripts in one list are linked with others in the same position in other lists. For example, we might re-arrange them like this

<u>location</u> $i=$	<u>age</u> $r=$	<u>time</u> $t=$	
1	1	1	(Subscripts are linked across rows here)
2	2	2	
3	3	3	
⋮	⋮	⋮	
N	R	λ	

Attached to the variable K these lists would mean that a person was in region ~~i_1 aged r_1 at t_1~~ 1 in age group 1 at time 1, in region ~~i_2 aged r_2 at t_2~~ 2 in age group 2 at time 2, and so on. The sex subscript does not similarly have to be listed n times because it is, except in rare cases, time-independent.

We could add other suitable subscripts to the three given above like social class $\cancel{c_1, c_2, c_3, \ldots, c_n}$ $c = 1, 2, 3, \ldots, m$ (Rees, 1972) or person type $b = 1, 2, 3,$ and 4 $\cancel{b_1, b_2, b_3, \ldots, b_4}$ (Wilson, 1972b), or use others like ethnic group E which behave in most respects like the sexes. But here we concentrate on aggregate populations (as in Rees and Wilson, 1973), and usually we will be concerned only with variables connected with

$$K^{ij}(t, t+T)$$

or the population in region i at time t who are in region j at time t+T.

To complete our definitions we need to represent in a convenient notation the events of birth and death. These take place at locations and are counted for regions. They represent transitions between a state of non-existence and existence (births) or between existence and non-existence (deaths). We can define a set of state labels for event and location:

$\beta(i) = \beta(1), \beta(2), \beta(3), \ldots, \beta(N)$

$\cancel{i_{1\beta}, i_{2\beta}, i_{3\beta}, \ldots, i_{n\beta}}$ for births and

$\cancel{i_{1\delta}, i_{2\delta}, i_{3\delta}, \ldots, i_{n\delta}}$ for deaths.

$\delta(i) = \delta(1), \delta(2), \delta(3), \ldots, \delta(N)$

These represent respectively "states of beginning life in region $1, 2, 3, \ldots, N$ $\cancel{i_1, i_2, i_3, \ldots, i_n}$" and "states of ending life in region $1, 2, 3, \ldots, N$." $\cancel{i_1, i_2, i_3, \ldots, i_n}$". The birth subscript will always start a locational subscript list and the death subscript will always end it. This is an improved version of the notation (i', j') used in Rees (1972) and has the advantage of distinguishing between births and deaths. An alternative (Rees, 1971) would be to place these subscripts at an intermediate position in the list at the time at which the event took place and fill the rest of the list with non-existence symbols, $\emptyset$,

e.g. $\emptyset, \emptyset, i, j$

at $\cancel{t_1, t_2, t_3, t_4}$ $t = 1, 2, 3, 4$

but this proved to be very awkward to handle.

3. The population accounts re-expressed

The population accounts described by Rees and Wilson (1973) can be re-expressed in this revised notation. The world is divided into two regions, region i of interest and R_i, the remainder or rest of the world, which is the world R less region i.* This is a completely closed system which can easily be generalized to n regions.**

The 16 population flows of Wilson and Rees (1973) are represented as follows:

	Old notation	Revised notation		Meaning
(1)	B^{i}	$K^{i\beta i}$	$K^{\beta(i)i}$	birth in i and survival in i
(2)	BM^{iR_i}	$K^{i\beta R_i}$	$K^{\beta(i)R_i}$	birth in i, migration to the rest of world and survival there
(3)	BD^{i}	$K^{i\beta i\delta}$	$K^{\beta(i)\delta(i)}$	birth in i and death in i
(4)	BMD^{iR_i}	$K^{i\beta R_i\delta}$	$K^{\beta(i)\delta(R_i)}$	birth in i, migration to the rest of world, death there
(5)	$BM^{R_i i}$	$K^{R_i\beta i}$	$K^{\beta(R_i)i}$	birth in the rest of world, migration to i and survival there
(6)	$BMD^{R_i i}$	$K^{R_i\beta i\delta}$	$K^{\beta(R_i)\delta(i)}$	birth in the rest of world, migration to i and death there
(7)	S^{i}	K^{ii}	K^{ii}	survival in i
(8)	M^{iR_i}	K^{iR_i}	K^{iR_i}	migration from i to the rest of the world
(9)	D^{i}	$K^{ii\delta}$	$K^{i\delta(i)}$	death in i
(10)	MD^{iR_i}	$K^{iR_i\delta}$	$K^{i\delta(R_i)}$	migration from i to the rest of the world and death there

* This introduces the principle that "rest of" regions are formed by subtracting the zone of interest i from the wider zone:

$$R_i = R - i$$

in geometric terms. Zones in the fixed boundaries are labelled with capital letters, variable zones with small case letter.

** Just to consider region i and j would not be completely general since the system would not be closed.

	Old notation	Revised notation	Meaning
(11)	$M^{R_i i}$	$\cancel{K^{R_i i}}$ $K^{R_i i}$	migration from the rest of the world to i and survival there
(12)	$MD^{R_i i}$	$\cancel{K^{R_i i_\delta}}$ $K^{R_i \delta(i)}$	migration from the rest of the world to i and death there
(13)	B^{R_i}	$\cancel{K^{R_{i\beta} R_i}}$ $K^{\beta(R_i) R_i}$	birth in the rest of the world and survival there
(14)	BD^{R_i}	$\cancel{K^{R_{i\beta} R_{i\delta}}}$ $K^{\beta(R_i)\delta(R_i)}$	birth in the rest of the world and death there
(15)	S^{R_i}	$\cancel{K^{R_i R_i}}$ $K^{R_i R_i}$	survival in the rest of the world
(16)	D^{R_i}	$\cancel{K^{R_i R_{i\delta}}}$ $K^{R_i \delta(R_i)}$	death in the rest of the world

Attached to each of these flows are the time subscripts t, t+T which indicate that the flows take place over that period.

These 16 flows can be re-arranged into an accounting table for the period t, t+T (Table 1). The principal advantage of the revised notation then becomes clear. Column and row sums can be formed algebraically using the subscripts. The summation asterisk * means summation over the relevant row or column elements. Row sums will be of zones $j = 1, \ldots, N, 1_\delta, \ldots, N_\delta$, column sums will be of zones $j = 1, \ldots, N, 1_\beta, \ldots, N_\beta$, and the grand total for the table will be over zones $j = 1, \ldots, N, 1_\beta, \ldots, N_\beta, 1_\delta, \ldots, N_\delta$. Table 1 can be converted into multi-regional form by replacing the R_i subscript by the appropriate summation over zones j. This is done in Table 2. It is usual to simplify complicated summations such as those in Table 2 by using asterisks. However, it should be kept in mind that asterisks in different positions in the subscript list imply that the summation has been performed over a different set of zones.

Table 1. The population accounts table in the revised notation

t \ t + T		Region i		Region R_i		Totals
		Alive at t + T	Died in t to t + T	Alive at t + T	Died in t to t + T	
Region i	Alive at t	K^{ii}	$K^{ii\delta}$	K^{iR_i}	$K^{iR_i\delta}$	K^{i*}
	Born in t to t+T	$K^{i_\beta i}$	$K^{i_\beta i_\delta}$	$K^{i_\beta R_i}$	$K^{i_\beta R_{i\delta}}$	$K^{i_\beta *}$
Region R_i	Alive at t	$K^{R_i i}$	$K^{R_i i\delta}$	$K^{R_i R_i}$	$K^{R_i R_{i\delta}}$	$K^{R_i *}$
	Born in t to t+T	$K^{R_{i\beta} i}$	$K^{R_{i\beta} i_\delta}$	$K^{R_{i\beta} R_i}$	$K^{R_{i\beta} R_{i\delta}}$	$K^{R_{i\beta}*}$
Totals		K^{*i}	$K^{*i\delta}$	K^{*R_i}	$K^{*R_{i\delta}}$	K^{**}

Notes

All the flows refer to period t to t+T.

4. The accounting equations re-expressed

The population of a region at the beginning of a period is given by

$$K^{i*}(t,t+T) = K^{ii}(t,t+T) + K^{ii\delta}(t,t+T) + \sum_{j\neq i}^{N} K^{ij}(t,t+T) + \sum_{j_\delta \neq i_\delta}^{N} K^{ij\delta}(t,t+T) \tag{1}$$

Population stocks are populations which have only one locational superscript explicit, the other has been summed over. Population stocks therefore can be re-written as

$K^{i*}(t)$

because they are invariant over the second time subscript. If we make the (t,t+T) subscripts implicit, equation (1) simplifies to

$$K^{i*}(t) = K^{ii} + K^{ii_\delta} + \sum_{j\neq i}^{N} K^{ij} + \sum_{j_\delta \neq i_\delta}^{N_\delta} K^{ij_\delta} \tag{2}$$

or even more simply to

$$K^{i*}(t) = \sum_{j=1}^{N} K^{ij} + \sum_{j_\delta = 1_\delta}^{N_\delta} K^{ij_\delta} \tag{3}$$

or in simplest form to

$$K^{i*}(t) = \sum_{j=1}^{N_\delta} K^{ij} \quad \text{where } j = 1,\ldots,N,\ 1_\delta,\ldots,N_\delta \tag{4}$$

The accounting equation for the population at the end of the period in region i becomes

$$K^{*i}(t+T) = K^{ii} + K^{i_\beta i} + \sum_{j\neq i}^{N} K^{ji} + \sum_{j_\beta \neq i_\beta}^{N_\beta} K^{j_\beta i} \tag{5}$$

which simplifies to

$$K^{*i}(t+T) = \sum_{j=1}^{N} K^{ji} + \sum_{j_\beta = 1_\beta}^{N_\beta} K^{j_\beta i} \tag{6}$$

or in simplest form

$$K^{*i}(t+T) = \sum_{j=1}^{N_\beta} K^{ji} \quad \text{where } j = 1, \ldots ,N,\ 1_\beta, \ldots ,N_\beta \tag{7}$$

The supporting equation for total deaths in region i is

$$K^{*i_\delta} = K^{ii_\delta} + K^{i_\beta i_\delta} + \sum_{j\neq i}^{N} K^{ji_\delta} + \sum_{j_\beta \neq i_\beta}^{N_\beta} K^{j_\beta i_\delta} \tag{8}$$

which simplifies to

$$K^{*i_\delta} = \sum_{j=1}^{N} K^{ji_\delta} + \sum_{j_\beta = 1_\beta}^{N_\beta} K^{j_\beta i_\delta} \tag{9}$$

In simplest form total deaths are given by

$$K^{*i_\delta} = \sum_{j=1}^{N_\beta} K^{ji} \quad \text{where } j=1,\ldots,N,1_\beta,\ldots,N_\beta \tag{10}$$

Table 2. The population accounts table for a single region recognizing many regions

		Region i		Region R_i , $k \in R_i$		Totals
		Alive at t+T	Died in t,t+T	Alive at t+T	Died in t,t+T	
Region i	Alive at t	K^{ii}	K^{ii_δ}	$\sum_{k \neq i}^{N} K^{ik}$ $k=1,\ldots,N$	$\sum_{k_\delta \neq i_\delta}^{N_\delta} K^{ik_\delta}$ $k_\delta=1_\delta,\ldots,N_\delta$	$\sum_{k}^{N_\delta} K^{ik}$ $k=1,\ldots,N,1_\delta,\ldots,N_\delta$
	Born in t,t+T	$K^{i_\beta i}$	$K^{i_\beta i_\delta}$	$\sum_{k \neq i}^{N} K^{i_\beta k}$ $k=1,\ldots,N$	$\sum_{k_\delta \neq i_\delta}^{N_\delta} K^{i_\beta k_\delta}$ $k_\delta=1_\delta,\ldots,N_\delta$	$\sum_{k}^{N_\delta} K^{i_\beta k}$ $k=1,\ldots,N,1_\delta,\ldots,N_\delta$
Region R_i $j \in R_i$	Alive at t	$\sum_{j \neq i}^{N} K^{ji}$ $j=1,\ldots,N$	$\sum_{j \neq i}^{N} K^{ji_\delta}$ $j=1,\ldots,N$	$\sum_{j \neq i}^{N} \sum_{k \neq i}^{N} K^{jk}$ $j=1,\ldots,N$ $k=1,\ldots,N$	$\sum_{j \neq i}^{N} \sum_{k_\delta \neq i_\delta}^{N_\delta} K^{jk_\delta}$ $j=1,\ldots,N$ $k_\delta=1_\delta,\ldots,N_\delta$	$\sum_{j \neq i}^{N} \sum_{k}^{N_\delta} K^{jk}$ $j=1,\ldots,N$ $k=1,\ldots,N,1_\delta,\ldots,N_\delta$
	Born in t,t+T	$\sum_{j_\beta \neq i_\beta}^{N_\beta} K^{j_\beta i}$ $j_\beta=1_\beta,\ldots,N_\beta$	$\sum_{j_\beta \neq i_\beta}^{N_\beta} K^{j_\beta i_\delta}$ $j_\beta=1_\beta,\ldots,N_\beta$	$\sum_{j_\beta \neq i_\beta}^{N_\beta} \sum_{k \neq i}^{N} K^{j_\beta k}$ $j_\beta=1_\beta,\ldots,N_\beta$ $k=1,\ldots,N$	$\sum_{j_\beta \neq i_\beta}^{N_\beta} \sum_{k_\delta \neq i_\delta}^{N_\delta} K^{j_\beta k_\delta}$ $j_\beta=1_\beta,\ldots,N_\beta$ $k_\delta=1_\delta,\ldots,N_\delta$	$\sum_{j_\beta \neq i_\beta}^{N_\beta} \sum_{k}^{N_\delta} K^{j_\beta k}$ $j_\beta=1_\beta,\ldots,N_\beta$ $k=1,\ldots,N,1_\delta,\ldots,N_\delta$
Totals		$\sum_{j}^{N_\beta} K^{ji}$ $j=1,\ldots,N,1_\beta,\ldots,N_\beta$	$\sum_{j}^{N_\beta} K^{ji_\delta}$ $j=1,\ldots,N,1_\beta,\ldots,N_\beta$	$\sum_{j}^{N_\beta} \sum_{k \neq i}^{N} K^{jk}$ $j=1,\ldots,N,1_\beta,\ldots,N_\beta$ $k=1,\ldots,N$	$\sum_{j}^{N_\beta} \sum_{k_\delta \neq i_\delta}^{N_\delta} K^{jk_\delta}$ $j=1,\ldots,N,1_\beta,\ldots,N_\beta$ $k_\delta=1_\delta,\ldots,N_\delta$	$\sum_{j}^{N_\beta} \sum_{k}^{N_\delta} K^{jk}$ $j=1,\ldots,N,1_\beta,\ldots,N_\beta$ $k=1,\ldots,N,1_\delta,\ldots,N_\delta$

The supporting equation for total births in region i is

$$K^{i_\beta *} = K^{i_\beta i} + K^{i_\beta i_\delta} + \sum_{j \neq i}^{N} K^{i_\beta j} + \sum_{j_\delta \neq i_\delta}^{N} K^{i_\beta j_\delta} \tag{11}$$

which simplifies to

$$K^{i_\beta *} = \sum_{j=1}^{N} K^{i_\beta j} + \sum_{j_\delta = 1_\delta}^{N_\delta} K^{i_\beta j_\delta} \tag{12}$$

and in simplest form to

$$K^{i_\beta *} = \sum_{j=1}^{N_\delta} K^{i_\beta j} \quad \text{where } j=1,\ldots,N,1_\delta,\ldots,N_\delta \tag{13}$$

In Rees and Wilson (1973) these accounting equations are used to fill in the accounts table once the small and unknown flows have been calculated. The translation of these procedures into the revised notation can be accomplished fairly easily and is not discussed in detail here. Rather we address some remaining difficulties with the location and state label lists which have to be cleared up.

5. A multiregional population accounts table

Table 3 sets out the population accounts table in multi-regional form (the equivalent table in Rees and Wilson is Figure 3). This table is an accounts table which focuses on the region: all flows which both start and end in a region are gathered together in clusters of four. The arrangement of rows and columns is, however, awkward from the computational point of view in the way that index numbers would be assigned location and state indexes:

location or state index of row	index number of row
1	1
1_β	2
2	3
2_β	4
⋮	⋮
N	2N-1
N_β	2N

Region at time t or t+T	Region at time t+T Alive in 1	Died in 1_δ	Alive in i	Died in i_δ	Alive in j	Died in j_δ	Alive in N	Died in N_δ	Totals
Alive in 1 Born in 1_β	K^{11} $K^{1_\beta 1}$	K^{11_δ} $K^{1_\beta 1_\delta}$	K^{1i} $K^{1_\beta i}$	K^{1i_δ} $K^{1_\beta i_\delta}$	K^{1j} $K^{1_\beta j}$	K^{1j_δ} $K^{1_\beta j_\delta}$	K^{1N} $K^{1_\beta N}$	K^{1N_δ} $K^{1_\beta N_\delta}$	K^{1*} $K^{1_\beta *}$
Alive in i Born in i_β	K^{i1} $K^{i_\beta 1}$	K^{i1_δ} $K^{i_\beta 1_\delta}$	K^{ii} $K^{i_\beta i}$	K^{ii_δ} $K^{i_\beta i_\delta}$	K^{ij} $K^{i_\beta j}$	K^{ij_δ} $K^{i_\beta j_\delta}$	K^{iN} $K^{i_\beta N}$	K^{iN_δ} $K^{i_\beta N_\delta}$	K^{i*} $K^{i_\beta *}$
Alive in j Born in j_β	K^{j1} $K^{j_\beta 1}$	K^{j1_δ} $K^{j_\beta 1_\delta}$	K^{ji} K^{ji}	K^{ji_δ} K^{ji_δ}	K^{jj} $K^{j_\beta j}$	K^{jj_δ} $K^{j_\beta j_\delta}$	K^{jN} $K^{j_\beta N}$	K^{jN_δ} $K^{j_\beta N_\delta}$	K^{j*} $K^{j_\beta *}$
Alive in N Born in N_β	K^{N1} $K^{N_\beta 1}$	K^{N1_δ} $K^{N_\beta 1_\delta}$	K^{Ni} $K^{N_\beta i}$	K^{Ni_δ} $K^{N_\beta i_\delta}$	K^{Nj} $K^{N_\beta j}$	K^{Nj_δ} $K^{N_\beta j_\delta}$	K^{NN} $K^{N_\beta N}$	K^{NN_δ} $K^{N_\beta N_\delta}$	K^{N*} $K^{N_\beta *}$
Totals	K^{*1}	K^{*1_δ}	K^{*i}	K^{*i_δ}	K^{*j}	K^{*j_δ}	K^{*N}	K^{*N_δ}	K^{**}

Table 3 The Population accounts table in Multi-Regional form

location or state index of row	index number of column
1	1
1_δ	2
2	3
2_δ	4
⋮	⋮
N	2N-1
N_δ	2N

The indexing of row and column locations and states can be more conveniently accomplished if the i_β and j_δ states follow the i and j states:

location or state index of row	index number of row	location or state index of column	index number of column
1	1	1	1
2	2	2	2
⋮	⋮	⋮	⋮
N	N	N	N
1_β	N+1	1_δ	N+1
2_β	N+2	2_δ	N+2
⋮	⋮	⋮	⋮
N_β	N+N	N_δ	N+N

This simply involves the transposition of rows and columns of Table 3 to form Table 4. The fundamental accounting relationships still hold but the focus of the table has switched from the region to the types of flows. All the survival and migration and survival terms are now grouped together in the top left hand portion of the table; all the birth and survival, and birth, migration and survival terms cluster in the bottom left; the death, and migration and death terms are concentrated in

TABLE 4 THE POPULATION ACCOUNTS TABLE RE-ARRANGED

		Region alive in, at time t+T							Region died in, time t to t+T							
		1	...	i	...	j	...	N	1_δ	...	i_δ	...	j_δ	...	N_δ	Totals
Region alive	1	K^{11}		K^{1i}		K^{1j}		K^{1N}	$K^{1 1_\delta}$		$K^{1 i_\delta}$		$K^{1 j_\delta}$		$K^{1 N_\delta}$	K^{1*}
in	⋮															⋮
at	i	K^{i1}		K^{ii}		K^{ij}		K^{iN}	$K^{i 1_\delta}$		$K^{i i_\delta}$		$K^{i j_\delta}$		$K^{i N_\delta}$	K^{i*}
time	⋮															⋮
t	j	K^{j1}		K^{ji}		K^{jj}		K^{jN}	$K^{j 1_\delta}$		$K^{j i_\delta}$		$K^{j j_\delta}$		$K^{j N_\delta}$	K^{j*}
	N	K^{N1}		K^{Ni}		K^{Nj}		K^{NN}	$K^{N 1_\delta}$		$K^{N i_\delta}$		$K^{N j_\delta}$		$K^{N N_\delta}$	K^{N*}
Region	1_β	$K^{1_\beta 1}$		$K^{1_\beta i}$		$K^{1_\beta j}$		$K^{1_\beta N}$	$K^{1_\beta 1_\delta}$		$K^{1_\beta i_\delta}$		$K^{1_\beta j_\delta}$		$K^{1_\beta N_\delta}$	$K^{1_\beta *}$
	⋮															⋮
born in	i_β	$K^{i_\beta 1}$		$K^{i_\beta i}$		$K^{i_\beta j}$		$K^{i_\beta N}$	$K^{i_\beta 1_\delta}$		$K^{i_\beta i_\delta}$		$K^{i_\beta j_\delta}$		$K^{i_\beta N_\delta}$	$K^{i_\beta *}$
time t to	⋮															⋮
t+T	j_β	$K^{j_\beta 1}$		$K^{j_\beta i}$		$K^{j_\beta j}$		$K^{j_\beta N}$	$K^{j_\beta 1_\delta}$		$K^{j_\beta i_\delta}$		$K^{j_\beta j_\delta}$		$K^{j_\beta N_\delta}$	$K^{j_\beta *}$
	N_β	$K^{N_\beta 1}$		$K^{N_\beta i}$		$K^{N_\beta j}$		$K^{N_\beta N}$	$K^{N_\beta 1_\delta}$		$K^{N_\beta i_\delta}$		$K^{N_\beta j_\delta}$		$K^{N_\beta N_\delta}$	$K^{N_\beta *}$
Totals		K^{*1}	...	K^{*i}	...	K^{*j}	...	K^{*N}	K^{*1_δ}	...	K^{*i_δ}	...	K^{*j_δ}	...	K^{*N_δ}	K^{**}

the upper right quadrant; and the birth and death, and birth, migration and death terms occupy the lower right quadrant. In matrix form Table 4 might be summarized as

	Alive at t+T	Died in t,t+T	Totals
Alive at t	$\underline{\underline{K}}^{ij}$	$\underline{\underline{K}}^{ij_\delta}$	$\underline{K}^{i*}$
Born in t,t+T	$\underline{\underline{K}}^{i_\beta j}$	$\underline{\underline{K}}^{i_\beta j_\delta}$	$\underline{K}^{i_\beta *}$
Totals	$\underline{K}^{*j}$	$\underline{K}^{*j_\delta}$	K^{**}

Table 5

All terms refer to the period (t,t+T)

where

$\underline{\underline{K}}^{ij}$ is the N*N survivorship and migration matrix

$\underline{\underline{K}}^{ij_\delta}$ is the N*N death and migration and death matrix

$\underline{\underline{K}}^{i_\beta j}$ is the N*N birth and survival, and birth, migration and survival matrix

$\underline{\underline{K}}^{i_\beta j_\delta}$ is the N*N birth and death, and birth, migration and death matrix

$\underline{K}^{i*}$ is the N 1 column vector of regional populations at time t

$\underline{K}^{*j}$ is the 1*N row vector of regional populations at time t+T

We might note parenthetically that five of the transition matrices possible in a birth-death-survival system are unrecognised. No transitions are allowed between died in t,t+T and alive at t+T (no resurrection), between alive at t and born in t,t+T (no rebirth), no death to death transitions (these are in heaven or hell presumably), no birth to birth transitions, and no death to birth transitions (re-incarnation).

These transitions are left to the theologians!

6. Rates and populations at risk

6.1 In Rees and Wilson (1973) it was argued that some four kinds of rates needed to be defined on the basis of the accounts:

(1) migration rates

(2) birth and migration rates

(3) total birth rates

(4) total death rates

Each rate calculation involved a numerator which was a population flow or population flow aggregate and a denominator which was the population at risk of being involved directly (or indirectly) in the population flow. This would include mothers at risk of giving birth to babies for example. That is,

model rate = population flow/population at risk

6.2 Populations at risk of birth and death

The notion of population at risk was explained in Rees and Wilson (1973) in case of total birth and total death rates. This notion is probably very close to the traditional population used in demographic rate calculations (Bogu , 1959) and to the $P_i(x)$ variable used to calculate an annual death rate by Rogers (1973, equations 1.4 and 1.5). In both the death rate and birth rate cases the populations at risk are made up of all persons likely to be involved in the event in a particular region, weighted by the length of time they spend there. In the revised K notation, population at risk of dying in region i is

$$\begin{array}{l}\text{population at risk}\\ \text{of dying in region } i\\ \text{in period } (t,t+T)\end{array} = \sum_{i=1}^{N_\beta} \sum_{j=1}^{N_\delta} {}^{i}\theta^{ij}(t,t+T)\ K^{ij}(t,t+T) \qquad (14)$$

where

${}^{i}\theta^{ij}(t,t+T)$ is the fraction of the period T spend in region i by people in the ij population flow (any of the population flows in the accounting table).

Note that a ${}^{i}\Theta^{kj}$ (t,t+T) is perfectly possible though normally it would be assumed to be <u>zero</u>.* More explicitly these modifying coefficients could be defined as

$$ {}^{i}\Theta^{jk}(t,t+T) = {}^{i}v^{jk}(t,t+T)/T \qquad (15)$$

where ${}^{i}v^{jk}(t,t+T)$ is the time spent by the K^{ij} (t,t+T) population flow in region i in the period t,t+T. This is divided by T the length of the period. The maximum value of any ${}^{i}\Theta^{jk}$ (t,t+T) will be <u>one</u> in the case of ${}^{i}\Theta^{ii}(t,t+T)$ where the persons involved in the flow haven't made any out-and-return migrations (unlike the author). The minimum value will be zero when no time is spent by a particular population flow in another region. The ${}^{i}\Theta^{jk}$ (t,t+T)'s are the aggregate population equivalents of Roger's $a_{ij}(x)$'s (Rogers, 1973 section 1.4.3.)

In the revised notation population at risk of dying would appear to be best notated as

$$\sum_{j=1}^{N_\beta}\sum_{k=1}^{N_\delta} {}^{i}K^{jk}(t,t+T)$$

but this in fact is a direct aggregate which simply counts up the number of people who were ever in region i in period (t,t+T). The population at risk of dying is a weighted function of the components and should probably be written

$$\Theta\Big(\sum_{j=1}^{N_\beta}\sum_{k=1}^{N_\delta} {}^{i}K^{jk}(t,t+T)\Big) = \sum_{j=1}^{N_\beta}\sum_{k=1}^{N_\delta} {}^{i}\Theta^{jk}(t,t+T)\, K^{ij}(t,t+T) \qquad (16)$$

where Θ is used as a modifying function on the left hand side.

---

* Consider, for example, that part of the author's lifeline between April 24/25, 1966 and April, 25/26, 1971, the last two census dates in the U.K. The author was recorded a survivor in the U.K. but in fact spent 4 out of the 5 years in the U.S.A.!

His

$$ {}^{(UK)}\Theta^{(UK)(UK)}(1966,1971) = 0.2 $$

but his

$$ {}^{(USA)}\Theta^{(UK)(UK)}(1966,1971) = 0.8 $$

The population at risk of giving birth would be

$$\Theta\left(\sum_{j=1}^{N}\sum_{k=1}^{N_\delta} {}^{i}K^{jk}\,(t,t+T)\right) = \sum_{j=1}^{N}\sum_{k=1}^{N_\delta} {}^{i}\Theta^{jk}\,(t,t+T)\;K^{ij}\,(t,t+T) \qquad (17)$$

where all birthflows are ignored.

Perhaps even more logical would be to place the at risk region subscript in an intermediate position in the locational list but not worry about supplying intermediate time subscripts or age subscripts (in the age disaggregated equivalent):

population at risk of dying might be represented as $\Theta(K^{jik}\,(t,t+T)$

It is clear from use of the revised notation that population at risk is dimensionally population, not population × time. It is measured in population units, not person-year units (as Rogers, 1973, equation 1.5 has it).

6.3 Populations at risk of migrating

It was argued in Rees and Wilson (1973) that time t population of the appropriate region is the population at risk of migration and survival, and that total births in the region of origin are the divisor need in the birth, migration and survival rate calculations. These are the only populations from which such population flows can emanate.

If migration is measured as a border crossing event, however, and counted like births and deaths (as are international migrations between countries) then a very different populations at risk may be involved. A number of other conceptual problems are involved in the use of migration as an event count data. These require solution. Is there a case, however, to use the same kind of population at risk for migration and survival as was used for births and deaths? This population at risk involved weighting population flows by average time exposed in the region of interest.

7. <u>The demographic model for a multi-regional, aggregate population system</u>

The matrix model of population projection originally stated by Leslie (1945) and developed by Keyfitz (1964a, 1964b, 1968) and by Rogers (1966, 1968, 1973) can be re-stated in the notation developed in this paper and in the accounts and models framework of Rees and Wilson (1973).

A matrix of transition rates can be defined which correspond to the flows table (Table 4) by dividing row elements by row totals to yield a matrix R which when transposed and multiplied by the vector of row sums yields a vector of column sums which include the new or end of period populations:

$$\underline{K}^{*i} = \underline{\underline{R}}\ \underline{K}^{i*} \tag{23}$$

or

$$\left[\begin{array}{c} K^{*1} \\ \vdots \\ K^{*i} \\ \vdots \\ K^{*j} \\ \vdots \\ K^{*N} \\ \hline K^{*1\delta} \\ \vdots \\ K^{*i\delta} \\ \vdots \\ K^{*j\delta} \\ \vdots \\ K^{*N\delta} \end{array}\right] = \left[\begin{array}{cccc|cccc}
\frac{K^{11}}{K^{1*}} \cdots & \frac{K^{i1}}{K^{i*}} \cdots & \frac{K^{j1}}{K^{j*}} \cdots & \frac{K^{N1}}{K^{N*}} & \frac{K^{1\beta 1}}{K^{1\beta *}} \cdots & \frac{K^{i\beta 1}}{K^{i\beta *}} \cdots & \frac{K^{j\beta 1}}{K^{j\beta *}} \cdots & \frac{K^{N\beta 1}}{K^{N\beta *}} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\frac{K^{1i}}{K^{1*}} \cdots & \frac{K^{ii}}{K^{i*}} \cdots & \frac{K^{ji}}{K^{j*}} \cdots & \frac{K^{Ni}}{K^{N*}} & \frac{K^{1\beta i}}{K^{1\beta *}} \cdots & \frac{K^{i\beta i}}{K^{i\beta *}} \cdots & \frac{K^{j\beta i}}{K^{j\beta *}} \cdots & \frac{K^{N\beta i}}{K^{N\beta *}} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\frac{K^{1j}}{K^{1*}} \cdots & \frac{K^{ij}}{K^{i*}} \cdots & \frac{K^{jj}}{K^{j*}} \cdots & \frac{K^{Nj}}{K^{N*}} & \frac{K^{1\beta j}}{K^{1\beta *}} \cdots & \frac{K^{i\beta j}}{K^{i\beta *}} \cdots & \frac{K^{j\beta j}}{K^{j\beta *}} \cdots & \frac{K^{N\beta j}}{K^{N\beta *}} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\frac{K^{1N}}{K^{1*}} \cdots & \frac{K^{iN}}{K^{i*}} \cdots & \frac{K^{jN}}{K^{j*}} \cdots & \frac{K^{NN}}{K^{N*}} & \frac{K^{1\beta N}}{K^{1\beta *}} \cdots & \frac{K^{i\beta N}}{K^{i\beta *}} \cdots & \frac{K^{j\beta N}}{K^{j\beta *}} \cdots & \frac{K^{N\beta N}}{K^{N\beta *}} \\
\hline
\frac{K^{11\delta}}{K^{1*}} \cdots & \frac{K^{i1\delta}}{K^{i*}} \cdots & \frac{K^{j1\delta}}{K^{j*}} \cdots & \frac{K^{N1\delta}}{K^{N*}} & \frac{K^{1\beta 1\delta}}{K^{1\beta *}} \cdots & \frac{K^{i\beta 1\delta}}{K^{i\beta *}} \cdots & \frac{K^{j\beta 1\delta}}{K^{j\beta *}} \cdots & \frac{K^{N\beta 1\delta}}{K^{N\beta *}} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\frac{K^{1i\delta}}{K^{1*}} \cdots & \frac{K^{ii\delta}}{K^{i*}} \cdots & \frac{K^{ji\delta}}{K^{j*}} \cdots & \frac{K^{Ni\delta}}{K^{N*}} & \frac{K^{1\beta i\delta}}{K^{1\beta *}} \cdots & \frac{K^{i\beta i\delta}}{K^{i\beta *}} \cdots & \frac{K^{j\beta i\delta}}{K^{j\beta *}} \cdots & \frac{K^{N\beta i\delta}}{K^{N\beta *}} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\frac{K^{1j\delta}}{K^{1*}} \cdots & \frac{K^{ij\delta}}{K^{i*}} \cdots & \frac{K^{jj\delta}}{K^{j*}} \cdots & \frac{K^{Nj\delta}}{K^{N*}} & \frac{K^{1\beta j\delta}}{K^{1\beta *}} \cdots & \frac{K^{i\beta j\delta}}{K^{i\beta *}} \cdots & \frac{K^{j\beta j\delta}}{K^{j\beta *}} \cdots & \frac{K^{N\beta j\delta}}{K^{N\beta *}} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\frac{K^{1N\delta}}{K^{1*}} \cdots & \frac{K^{iN\delta}}{K^{i*}} \cdots & \frac{K^{jN\delta}}{K^{j*}} \cdots & \frac{K^{NN\delta}}{K^{N*}} & \frac{K^{1\beta N\delta}}{K^{1\beta *}} \cdots & \frac{K^{i\beta N\delta}}{K^{i\beta *}} \cdots & \frac{K^{j\beta N\delta}}{K^{j\beta *}} \cdots & \frac{K^{N\beta N\delta}}{K^{N\beta *}}
\end{array}\right] \times \left[\begin{array}{c} K^{1*} \\ \vdots \\ K^{i*} \\ \vdots \\ K^{j*} \\ \vdots \\ K^{N*} \\ \hline K^{1\beta *} \\ \vdots \\ K^{i\beta *} \\ \vdots \\ K^{j\beta *} \\ \vdots \\ K^{N\beta *} \end{array}\right] \tag{24}$$

This is the population projection consistent with the demographic accounts outlined in Rees and Wilson (1973) and in this paper. To use the model, however, the account building procedures developed in Rees and Wilson (1973) have to be used. The rate definitions therefore change and the row and column vectors in the model simplify to old and new population vectors. Further explorations of the matrix version of the population projection model stated here are obviously suggested, and a statement of the relation of this model and earlier ones ~~would be valuable.~~ is made in Rees and Wilson (1972).

3. Conclusion

This paper has introduced a revised notation in which the accounts and models of Rees and Wilson (1973) can be expressed. While no new concepts have been introduced old ones have been clarified.

The major disadvantage of the revised notation is that it is probably harder to grasp the fundamentals of the accounting system through subscripts rather than through variable letters.

The major advantages of the revised notation are as follows:

(1) The same entity, a person, is similarly labelled, no matter what the situation. The K variable means population in any context. The previous notation made it appear that we were adding apples to oranges.

(2) The history of a set of persons in the same population flow is contained in a subscript list. This has been previously demonstrated to be a very effective theory building tool (Wilson, 1972).

(3) Consistency checks on any equation are very easily accomplished by comparing subscripts in the same position in any variable list.

(4) The states of "birth in the period" and "death in the period" are distinguished. They were not in Rees (1972).

(5) The revised notation meets the objections of mathematicians who regard double or triple letter variables as introducing the sins of computerese into mathematics. Algebraic confusion is avoided.

(6) This notation should aid the linking of the accounts and models of Rees and Wilson (1973) with the work of Keyfitz and Rogers.

References

Batty, M. (1970). "An activity allocation model for the Nottinghamshire - Derbyshire sub-region". Regional Studies, 4, pp.307-332.

Bogue, D.J. (1959). The principles of demography. John Wiley, New York.

Cripps, E.L. and Foot, D.H.S. (1969). "A land-use model for sub-regional planning". Regional Studies, 3, pp.243-268.

Keyfitz, N. (1964a). "Matrix multiplication as a technique of population analysis". The Millbank Memorial Fund Quarterly, 42, pp.68-84.

Keyfitz, N. (1964b). "The population projection as a matrix operator". Demography, 1, pp.57-73.

Keyfitz, N. (1968). Introduction to the mathematics of population. Addison-Wesley, Reading Massachusetts.

Leslie, P.H. (1945). "On the use of matrices in certain population mathematics". Biometrika, 33, pp.185-212.

Rees, P.H. (1971). "Constructing a set of population accounts for a region". Working Note, Department of Geography, University of Leeds.

Rees, P.H. (1972). "The distribution of social groups within cities: accounts and models". Patterns and processes in urban and regional systems. Ed. A.G. Wilson. Pion, London.

*

Rees, P.H. and Wilson, A.G. (1973). "Accounts and models for spatial demographic analysis I: aggregate populations". Environment and Planning, 5, No.1, forthcoming.

Rogers, A. (1966). "Matrix methods of population analysis". Journal of the American Institute of Planners, 30, pp.40-44.

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

Rogers, A. (1973). "The mathematics of multi-regional demographic growth". Environment and Planning, 5, No.1, forthcoming.

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

Wilson, A.G. (1972a). "Multi-regional models of population structure and some implications for a dynamic-residential location model". Patterns and processes in urban and regional analysis. Ed. A.G. Wilson. Pion, London.

Wilson, A.G. (1972b). "Principles for building a general urban and regional model". Working Paper 21, Department of Geography, University of Leeds.

INSERT * Rees, P.H. and Wilson, A.G. (1972) A matrix formulation of accounts and models for spatial demographic analysis. Working Paper in preparation, Department of Geography, University of Leeds.