Environment and Planning, 1973, volume 5, pages 61-90

Accounts and models for spatial demographic analysis I: aggregate population

P H Rees, A G Wilson
Department of Geography, University of Leeds, Leeds, England
Received 21 September 1972

Abstract. This paper presents accounts and models which can be used for spatial demographic analysis. It is shown that in any time period there are sixteen different kinds of demographic flows for a two region system. These flows can be represented in an accounting table. Some of the minor flows, involving double or triple demographic events such as a birth followed by a migration, cannot be measured directly from data, and an iterative scheme is offered for the calculation of these flows. In this way a complete set of accounts can be built up from existing data. Two principles of model building using this set of accounts are then discussed. The first uses a simple definition of demographic rates, while the second uses a definition of rates which employs a concept of 'at-risk' populations. Procedures are given for calculating these populations. The account building principles are illustrated with a worked example for a three region system: the West Riding of Yorkshire, the rest of England and Wales, and the rest of the world. The results presented in this paper are for aggregate populations. In a subsequent paper equivalent results will be presented for age/sex disaggregated populations.

1 Introduction

Demography is concerned with population structure and change through processes of birth, survival, and migration. Usually demographers study the population of a single area or region with an 'outside world' recognised for migration purposes. One of the concerns of human geography is the spatial distribution of population, and the spatial variation of population structure and rates of change. It is necessary for adequate analysis in both disciplines, therefore, that multiregional demographic models are developed. Given such a problem it is useful to connect the model building task to the concept of population accounts. This ensures that proper connections are made between different parts of the model, both in relation to time periods and in relation to regions.

Multiregional demographic models have been developed by Rogers (1966, 1968) and generalised by Wilson (1972). Associated explorations of population accounting concepts have been made by Rees (1972a). In this paper we try to draw the two threads together, and our aim is to present a multiregional model for spatial demographic analysis, which is consistent with an underlying set of population accounts. We shall pay particular attention to the problem of formulating the model so that it can be used with the kind of census data that is usually available.

The models to be discussed are based on notions of 'rates'—births per thousand mothers of a certain age in a certain region, and so on. The accounts are based on population flows—total births in a region, migrants in an age group from one region to another in a certain time period, and so on. We begin, however, by outlining the accounting principles, and showing how models can be constructed from the fundamental accounting equations. In this paper we illustrate the basic principles of our approach using aggregate populations only—that is, with no age or sex categories. In a second forthcoming paper (Wilson and Rees, 1972a) we apply the same methods to age and sex disaggregated populations. In section 2 we introduce the accounting concepts, in section 3 we indicate how missing cells in the accounts tables can be filled, in section 4 we indicate how models can be developed, and in section 5 a worked example is presented.

2 Basic accounting relationships

Accounting consists of being able to specify the life history of the population units under consideration, whether these be persons or new decimal pence. In financial accounting every unit must be accounted for. In this section we specify a conceptual framework which would enable us in principle to do this for persons, though we shall see later that we are forced to make some approximations in practice.

For a single region, accounting relations are usually stated in terms of a single component of change equation for a time period, say from time t to time $t+T$. We might have

$$w(t+T) = w(t)-D(t,t+T)+B(t,t+T)+I(t,t+T)-O(t,t+T) , \quad (1)$$

where
$w(t)$ is the population at time t,
$D(t,t+T)$ are the deaths in period t to $t+T$,
$B(t,t+T)$ are the births in period t to $t+T$,
$I(t,t+T)$ is the inmigration in period t to $t+T$[1],
$O(t,t+T)$ is the outmigration in period t to $t+T$.
When the time period involved is clear from the content we shall drop the designation $(t,t+T)$ on flow variables for simplicity.

As noted earlier, our main emphasis is to be the development of a multiregional model. Thus, at the outset, we develop a multiregional framework. Suppose we have a system of N regions, $i = 1, 2, ..., N$. (Typically N itself will be the 'rest of the world'.) Then we shall specify accounting equations for each region i. It will also be useful to designate the set of regions, $R(i)$, as all regions excluding i. Then, as far as region i is concerned relative to $R(i)$, there are sixteen types of demographic event as depicted in figure 1.

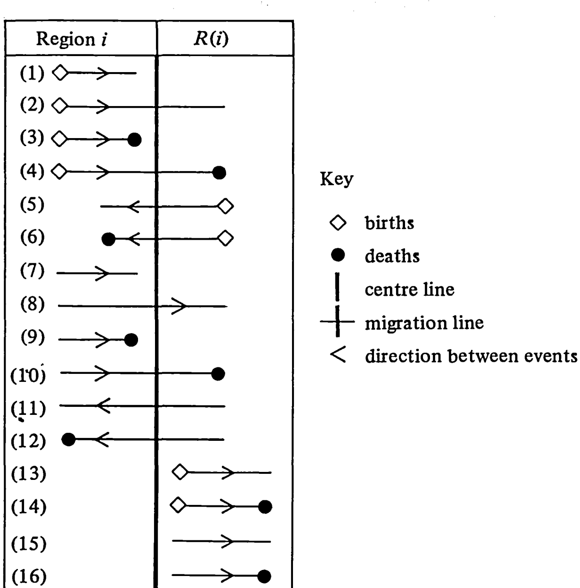

Figure 1. Types of demographic event for a two-region system.

[1] Unless otherwise stated we are using the census definition of migration, which means 'new address' in the form of change of region at the end of the period irrespective of possible intermediate moves. This also implies survival to the end of the period.

The types are as follows (all events in period t to $t+T$):
(1) Birth in region i and survival in region i,
(2) Birth in region i, migration to $R(i)$, and survival in $R(i)$,
(3) Birth in region i and death in region i,
(4) Birth in region i, migration to $R(i)$, and death in $R(i)$,
(5) Birth in $R(i)$, migration to region i, and survival in region i,
(6) Birth in $R(i)$, migration to region i, and death in region i,
(7) Survival in region i,
(8) Migration to $R(i)$ and survival in $R(i)$,
(9) Death in region i,
(10) Migration to $R(i)$ and death in $R(i)$,
(11) Migration to region i and survival in region i,
(12) Migration to region i and death in region i,
(13) Birth in $R(i)$ and survival in $R(i)$,
(14) Birth in $R(i)$ and death in $R(i)$,
(15) Survival in $R(i)$,
(16) Death in $R(i)$.

To deal with the multiregional case more formally, we need to extend the notation introduced after equation (1). We do this as follows: let
$w^i(t)$ be the population in region i at time t[2];
D^i be the deaths in region i during the period $(t, t+T)$ of people present in region i at the beginning of the period[3];
MD^{ij} be the deaths in region j of people who have migrated from region i[4];
B^i be the births in region i, resident in region i at end of period;
BD^i be the births in region i, death in region i;
BM^{ij} be the births in region i, plus migration to region j from i and survival in region j;
BMD^{ij} be the births in region i, migration from region i to j, and death in region j;
M^{ij} be the migration from region i to region j;
S^i be survival in region i.

The sequences of letters used in some variable names represent a sequence of demographic events. Then, if we consider interactions between i and $R(i)$, we can represent each of the sixteen flows of figure 1 in this notation as follows:

(1)	B^i	(5)	$BM^{R(i)i}$	(9)	D^i	(13)	$B^{R(i)}$
(2)	$BM^{iR(i)}$	(6)	$BMD^{R(i)i}$	(10)	$MD^{iR(i)}$	(14)	$BD^{R(i)}$
(3)	BD^i	(7)	S^i	(11)	$M^{R(i)i}$	(15)	$S^{R(i)}$
(4)	$BMD^{iR(i)}$	(8)	$M^{iR(i)}$	(12)	$MD^{R(i)i}$	(16)	$D^{R(i)}$.

Following Stone (1971), as developed in Rees (1972a)[5] we can incorporate these flows in the accounting framework shown in figure 2, where we have introduced:
TB^i as the total observed births in region i in the period,
TD^i as the total observed deaths in region i in the period,
and similarly $TB^{R(i)}$, and $TD^{R(i)}$ for region $R(i)$.

The first row of the table specifies what happens to the population in region i at time t over the time period $(t, t+T)$. Some members of the original population

[2] This is for any region i of course. Also, $w^{R(i)}(t)$ would be the population of $R(i)$ at t.
[3] t to $t+T$ is implied for all following flow definitions.
[4] Note that *MD* here, and *BD, BM, BMD, TB* and *TD* below represent single variables.
[5] The accounting table presented in figure 2 is a transformation of the single period, closed population accounting table used by Stone. Stone's table is a single-region table in which all other regions are considered to be 'in the outside world', births are considered to originate in the outside world, and deaths to terminate in the outside world. Figure 2 is therefore a multiregion expansion of Stone's table.

survived *in situ* (S^i), some died *in situ* (D^i), some migrated to other regions and survived there ($M^{iR(i)}$), and some migrated to other regions and died there ($MD^{iR(i)}$). These terms all sum to the population of region i at time t. The third row specifies what happens to the population of the rest of the world $R(i)$ in the same way. The second and fourth rows tell us what happened to the babies born respectively in region i (TB^i) and region $R(i)$ ($TB^{R(i)}$) over the period.

The first and third columns of the table tell us about the origins of the populations of regions i and $R(i)$, and the second and fourth columns tell us about the origins of the persons who have died. The columns inform us about where the populations came from in the period $(t, t+T)$.

The basic accounting relationships for region i can then be obtained by summing the elements of the first column:

$$w^i(t+T) = S^i + B^i + M^{R(i)i} + BM^{R(i)i} \,. \tag{2}$$

The shaded cells in figure 2 denote flows which are usually available in historical census data or registration records[6]. Thus equation (2) involves two flows, S^i and B^i, which are not available directly. S^i can be obtained from the accounting relationship implied by the first row as

$$S^i = W^i(t) - D^i - M^{iR(i)} - MD^{iR(i)} \,. \tag{3}$$

D^i, in turn, can be obtained by the relationship in the second column:

$$D^i = TD^i - BD^i - MD^{R(i)i} - BMD^{R(i)i} \,. \tag{4}$$

Location at time t	State at time t	Location at time $t+T$. Region i, State at time $t+T$: alive at $t+T$ (survivors)	Region i: died in t to $t+T$	Region $R(i)$: alive at $t+T$ (survivors)	Region $R(i)$: died in t to $t+T$	Totals
Region i	alive at t	S^i	D^i	$M^{iR(i)}$	$MD^{iR(i)}$	$w^i(t)$
	born in t to $t+T$	B^i	BD^i	$BM^{iR(i)}$	$BMD^{iR(i)}$	TB^i
Region $R(i)$	alive at t	$M^{R(i)i}$	$MD^{R(i)i}$	$S^{R(i)}$	$D^{R(i)}$	$w^{R(i)}(t)$
	born in t to $t+T$	$BM^{R(i)i}$	$BMD^{R(i)i}$	$B^{R(i)}$	$BD^{R(i)}$	$TB^{R(i)}$
	Totals	$w^i(t+T)$	TD^i	$w^{R(i)}(t+T)$	$TD^{R(i)}$	

Figure 2. Accounting framework for two regions.

(6) In figure 2, and below, we have assumed that S^i is not directly available. In fact, as pointed out to us by Andrei Rogers, it can sometimes be obtained directly as 'persons not moving' from Census *migration* tables. In circumstances where this is the case, alternative estimation procedures to the ones outlined below are possible. They will be explored in Wilson and Rees (1973).

B^i can be obtained by the relationship in the second row:

$$B^i = TB^i - BD^i - BM^{iR(i)} - BMD^{iR(i)} \, . \tag{5}$$

Equations (3), (4) and (5) still involve the quantities BD^i, $MD^{R(i)i}$, $MD^{iR(i)}$, $BMD^{R(i)i}$, and $BMD^{iR(i)}$ which are usually unknown. However, later we shall show how these quantities can be estimated from data with the help of an assumption. Thus we can substitute from equation (4) into (3), and from equations (3) and (5) into (2), to obtain

$$\begin{aligned} w^i(t+T) = {} & w^i(t) - (TD^i - BD^i - MD^{R(i)i} - BMD^{R(i)i}) \\ & + (TB^i - BD^i - BM^{iR(i)} - BMD^{iR(i)}) + BM^{R(i)i} \\ & + M^{R(i)i} \\ & - (M^{iR(i)} + MD^{iR(i)}) \, , \end{aligned} \tag{6}$$

where the terms on successive lines are equivalent to those in equation (1). We now have the correct account-based specification of deaths, births, inmigration and outmigration. The equation can be written out in words as

new population = old population − (deaths in old population)
+ (surviving non-migrating births, and surviving inmigrating infants)
+ (surviving inmigrants)
− (outmigrants) . (7)

Since this accounting relationship takes note of availability of information, it is the one we shall usually employ. Others could, however, be derived from equation (2) in a similar manner as appropriate.

We can summarise our argument by saying that for region i, four accounting equations are implied by the accounts of figure 2—the two rows and the two columns for region i. The first column gives the new population and would be the basis of

Region at time t		Region at time $t+T$												Totals
		1			i			j			N			
		alive	died		alive	died		alive	died		alive	died		
1	alive	S^1	D^1		M^{1i}	MD^{1i}		M^{1j}	MD^{1j}		M^{1N}	MD^{1N}		$w^1(t)$
	to be born	B^1	BD^1		BM^{1i}	BMD^{1i}		BM^{1j}	BMD^{1j}		BM^{1N}	BMD^{1N}		TB^1
i	alive	M^{i1}	MD^{i1}		S^i	D^i		M^{ij}	MD^{ij}		M^{iN}	MD^{iN}		$w^i(t)$
	to be born	BM^{i1}	BMD^{i1}		B^i	BD^i		BM^{ij}	BMD^{ij}		BM^{iN}	BMD^{iN}		TB^i
j	alive	M^{j1}	MD^{j1}		M^{ji}	MD^{ji}		S^j	D^j		M^{jN}	MD^{jN}		$w^j(t)$
	to be born	BM^{j1}	BMD^{j1}		BM^{ji}	BMD^{ji}		B^j	BD^j		BM^{jN}	BMD^{jN}		TB^j
N	alive	M^{N1}	MD^{N1}		M^{Ni}	MD^{Ni}		M^{Nj}	MD^{Nj}		S^N	D^N		$w^N(t)$
	to be born	BM^{N1}	BMD^{N1}		BM^{Ni}	BMD^{Ni}		BM^{Nj}	BMD^{Nj}		B^N	BD^N		TB^N
Totals		$w^1(t+T)$	TD^1		$w^i(t+T)$	TD^i		$w^j(t+T)$	TD^j		$w^N(t+T)$	TD^N		

Figure 3. Accounting framework for N regions.

any model estimate. We use the second column to estimate D^i, and the first and second rows to estimate S^i and B^i respectively.

The next step is to show how the accounting system shown in figure 2 can be extended to show all N regions explicitly. $R(i)$ is split into the individual regions to give figure 3.

3 Estimation of the unknown flows from known flows

In the previous section we showed how to estimate S^i, D^i, and B^i in equations (3), (4), and (5) using the accounting relationships. We now proceed to a more detailed discussion and show how the quantities BD^i, $BD^{R(i)}$, $MD^{R(i)i}$, $MD^{iR(i)}$, $BMD^{R(i)i}$, and $BMD^{iR(i)}$ are estimated.

In relation to the shaded cells of figure 2 let us assume we have satisfactorily estimated, from items in census and registration documents, the following flows and event totals for region i:

$M^{iR(i)}$, $M^{R(i)i}$, $BM^{iR(i)}$, $BM^{R(i)i}$, TB^i, $TB^{R(i)}$, TD^i, and $TD^{R(i)}$.[7]

We estimate S^i and B^i using equations (3) and (5) as indicated earlier. These equations require estimation of D^i and $MD^{iR(i)}$, and BD^i and $BMD^{iR(i)}$. The term D^i is given by equation (4), which requires the estimation of BD^i. The other terms occur in the corresponding equations for $R(i)$.

The general principle adopted in estimating the unknown flows is that persons involved in the corresponding events experienced the event rates of the region in which the event took place. For example, we assume that persons who outmigrate and die experience the death rate of the region in which they die, not that of the region they were in at the beginning of the time period. This is a reasonable first approximation which takes into account the immediate environmental effects on death rates, such as exposure to disease, environmental pollution, health care facilities, nature of climate, and so on[8].

To proceed we need death rates for i and $R(i)$. We shall discuss later two ways of estimating these; the better method utilises information on total deaths, and for this reason we shall call it the total death rate, written td^i for region i, $td^{R(i)}$ for region $R(i)$. The other rate will be called simply the death rate, and written as d^i or $d^{R(i)}$ as appropriate. It must be remembered that the adjective 'total' in the first definition refers to the method of calculation; it does not imply that, in principle, some deaths are omitted in the second definition. In the rest of this section we shall write x^i or $x^{R(i)}$ for the death rate, and later substitute td^i (or $td^{R(i)}$) or d^i (or $d^{R(i)}$) as required. We can now proceed to estimate, in turn, $MD^{iR(i)}$, $MD^{R(i)i}$, $BMD^{iR(i)}$, $BMD^{R(i)i}$, BD^i, and $BD^{R(i)}$.

Using our general principle, we have*

$$MD^{iR(i)} = x^{R(i)}(M^{iR(i)} + MD^{iR(i)}) , \tag{8}$$

(7) The methods for doing this estimation are explained in another paper (Rees, 1972b).

(8) Alternative assumptions are possible. In theory, perhaps, information on death rates in all regions of previous residences, weighted by length of residence, would be appropriate, but such an analysis would require a detailed longitudinal study of population cohorts and is obviously infeasible.

* *Note added in proof.* Further research has suggested that this argument can be improved. The principle is unaffected, but any reader who wishes to use this procedure should incorporate the following modifications. The right hand side of equation (8) should be $x^{R(i)}(0\cdot 5M^{iR(i)} + 0\cdot 25MD^{iR(i)})$, and the expression $x^{R(i)}/(1-x^{R(i)})$ in equations (9) and (11) then becomes $0\cdot 5x^{R(i)}/(1-0\cdot 25x^{R(i)})$. There are corresponding changes in equations (10) and (12). Because of a similar argument, there are changes in equations (13) and (14) which have the effect that x^i in equation (15) becomes $0\cdot 5x^i/(1+0\cdot 25x^i)$. Analogous changes should be made in equations containing similar terms in the rest of the paper.

since $M^{iR(i)} + MD^{iR(i)}$ is the total outmigration from i to $R(i)$, and we apply the $R(i)$ death rate. This gives

$$MD^{iR(i)} = \frac{x^{R(i)}}{1-x^{R(i)}} M^{iR(i)} \,. \tag{9}$$

By a similar argument

$$MD^{R(i)i} = \frac{x^{i}}{1-x^{i}} M^{R(i)i} \,, \tag{10}$$

$$BMD^{iR(i)} = \frac{x^{R(i)}}{1-x^{R(i)}} BM^{iR(i)} \,, \tag{11}$$

and

$$BMD^{R(i)i} = \frac{x^{i}}{1-x^{i}} BM^{R(i)i} \,. \tag{12}$$

The term BD^i could actually be cancelled from equation (6), but in disaggregated versions it reappears. Also it makes interpretation of the relevant equations easier, and so we retain the term and calculate it as follows. By rearranging equation (5), we get

$$B^i + BD^i = TB^i - BM^{iR(i)} - BMD^{iR(i)} \,, \tag{13}$$

and, using our principle,

$$BD^i = x^i(B^i + BD^i) \,, \tag{14}$$

so that, using equation (13),

$$BD^i = x^i(TB^i - BM^{iR(i)} - BMD^{iR(i)}) \,. \tag{15}$$

Similarly

$$BD^{R(i)} = x^{R(i)}(TB^{R(i)} - BM^{R(i)i} - BMD^{R(i)i}) \,. \tag{16}$$

B^i can now be calculated from equation (5) using known flows, BD^i from equation (15), and $BMD^{R(i)i}$ from equation (12). There is a similar procedure for $B^{R(i)}$. D^i and S^i can be calculated from equations (4) and (3) respectively, and $D^{R(i)}$ and $S^{R(i)}$ similarly[9].

So far we have assumed that $BM^{iR(i)}$ and $BM^{R(i)i}$ can be obtained directly. In many cases, however, this is not the case. Hence, we next give a procedure for estimating these flows in the case where they are not available from data.

We assume that babies migrate at the same rate as the rest of the population, so that

$$BM^{iR(i)} = \frac{M^{iR(i)}}{w^i(t)} TB^i \,, \tag{17}$$

with a similar equation for $BM^{R(i)i}$.

Usually, of course, we shall be dealing with age disaggregated populations, and the estimates can be improved by multiplying total births by the migration rate of the parental age group weighted by the proportion of births to that age group of parents.

[9] As an alternative, we could have used equation (4) as a basis for estimating BD^i. We might say

$$D^i = x^i(S^i + D^i) = x^i[w^i(t) - M^{iR(i)} - MD^{iR(i)}] \,,$$

using equation (3). So, rearranging equation (4),

$$BD^i = TD^i - x^i[w^i(t) - M^{iR(i)} - MD^{iR(i)}] - MD^{R(i)i} - BMD^{R(i)i} \,.$$

We consider this problem in our second paper which deals with the disaggregated case (Wilson and Rees, 1972a).

In this section the methods for filling in the blanks in the accounts table have been developed for aggregate populations in a two region context. This simplification is carried through into the next section which develops the principles of model building. However, the reader will realise that there is no loss of generality here. It is a simple task to replace $R(i)$ by j, for several j, and to make the appropriate modifications in the equations presented.

4 Principles of model building

4.1 *Introduction*

The task of a demographic model is to estimate a 'new' population, say $w^i(t+T)$, from an 'old' one, say $w^i(t)$. We discussed the traditional procedure in relation to equation (1), and then in equation (2) showed how to do this correctly in relation to the basic accounts presented as figure 2. Equation (2) involves two terms which are not directly available, S^i and B^i, and we showed how we could use the other three accounting equations for region i to estimate in turn S^i (in terms of D^i), D^i, and B^i in equations (3), (4), and (5). In section 3 we gave a procedure for estimating the minor demographic flows, and this procedure involved an assumption about death rates. Then provided we can make a reasonable assumption about death rates, we can use the analysis thus far to put together a complete set of historical accounts for i and $R(i)$ (or for the whole set of N regions).

If we are to make projections, however, the basic 'new population' equation, such as (2), must be expressed in terms of rates. These rates can then be projected, and the model used to make a population projection. The problem of defining these rates adequately is a difficult one, and so we proceed in two stages with what we call simple base population rates, and corrected at risk population rates, and we discuss the models generated by each.

4.2 *Model using simple base population rates*

The correct way to define event rates using the figure 2 accounts is to divide particular elements by appropriate row or column totals. Thus we might define, with little difficulty,

$$s^i = \frac{S^i}{w^i(t)} \tag{18}$$

as a survival rate, and

$$m^{R(i)i} = \frac{M^{R(i)i}}{w^{R(i)}(t)} \tag{19}$$

as an inmigration rate.

The birth terms B^i and $BM^{R(i)i}$ in equation (2) present more problems. The obvious divisors are TB^i and $TB^{R(i)}$ respectively. We shall return to this possibility later, but meanwhile we assume that we can use $w^i(t)$ and $w^{R(i)}(t)$ as divisors, and define rates

$$b^i = \frac{B^i}{w^i(t)} , \tag{20}$$

and

$$bm^{R(i)i} = \frac{BM^{R(i)i}}{w^{R(i)}(t)} . \tag{21}$$

By using equations (18)–(21), equation (2) can now be written

$$w^i(t+T) = s^i w^i(t) + b^i w^i(t) + m^{R(i)i} w^{R(i)}(t) + bm^{R(i)i} w^{R(i)}(t) . \tag{22}$$

Then if the rates s^i, b^i, $m^{R(i)i}$, and $bm^{R(i)i}$ could be projected and estimated, say for the period $t+T$ to $t+2T$, $w^i(t+2T)$ could be estimated.

This analysis illustrates the principles of demographic model building in the simplest possible way. However, we have already seen that such terms as S^i and B^i, and hence s^i and b^i, cannot be directly estimated, and that we must resort to the basic accounts and replace equation (2) by something like equation (6). Further, we should emphasise that when attempts are made to measure such terms as S^i and B^i without an accounting basis, it is all too easy to omit some of the small flows.

Equation (6) could be written in the form

$$w^i(t+T) = w^i(t) - D^i + B^i + BM^{R(i)i} + M^{R(i)i} - M^{iR(i)} - MD^{iR(i)} , \tag{23}$$

and we could define

$$d^i = \frac{D^i}{w^i(t)} = \frac{TD^i - BD^i - MD^{R(i)i} - BMD^{R(i)i}}{w^i(t)} \tag{24}$$

as a simple base population death rate,

$$m^{iR(i)} = \frac{M^{iR(i)}}{w^i(t)} \tag{25}$$

as an outmigration rate, and b^i, $bm^{R(i)i}$, and $m^{R(i)i}$ as in equations (20), (21), and (19), though with b^i now being alternatively written as

$$b^i = \frac{TB^i - BD^i - BM^{iR(i)} - BMD^{iR(i)}}{w^i(t)} . \tag{26}$$

Using equation (1), we can then define

$$md^{iR(i)} = \frac{d^{R(i)}}{1-d^{R(i)}} m^{iR(i)} , \tag{27}$$

where we have replaced $x^{R(i)}$ by the death rate being used here, $d^{R(i)}$ [defined in the $R(i)$ analogue of equation (24)]. Equation (23) can now be written

$$\begin{aligned} w^i(t+T) = {} & w^i(t) - d^i w^i(t) + b^i w^i(t) + bm^{R(i)i} w^{R(i)}(t) + m^{R(i)i} w^{R(i)}(t) \\ & - m^{iR(i)} w^i(t) - md^{iR(i)} w^i(t) . \end{aligned} \tag{28}$$

Then provided we can project the rates d^i, b^i, $bm^{R(i)i}$, $m^{R(i)i}$, $m^{iR(i)}$ [and calculate $md^{iR(i)}$ from equation (27)], equation (28) can be used as a model equation for projections. The actual procedure is complicated by the fact that the minor flows in the death rate definitions themselves depend on the death rate through the equations of section 3, and so all the equations must be solved iteratively. The procedure is summarised in the next subsection.

Of course, the models presented here as equations (22) and (28) are related, since

$$s^i = 1 - d^i - m^{iR(i)} - md^{iR(i)} . \tag{29}$$

However, the calculation procedure associated with equation (28) has to be used in relation to known data, but a survival rate could be calculated from equation (29) and used in equation (22) if desired.

4.3 *Summary of model building procedure using simple rates*

The procedure can be summarised in the following steps:

1 Estimate the following from census and registration records: $w^i(t)$, $w^{R(i)}(t)$, $M^{iR(i)}$, $M^{R(i)i}$, TD^i, $TD^{R(i)}$, $BM^{iR(i)}$, $BM^{R(i)i}$, TB^i, and $TB^{R(i)}$.

2 Estimate the death rate d^i from equation (23), and $d^{R(i)}$ from an equivalent equation, *initially* setting BD^i, $MD^{R(i)i}$, $BMD^{R(i)i}$, and the equivalent $R(i)$ terms to zero.

3 Calculate $MD^{iR(i)}$, $MD^{R(i)i}$, $BMD^{R(i)i}$, $BMD^{iR(i)}$, BD^i, and $BD^{R(i)}$ from equations (9)–(12), (15), and (16) respectively with $x^i = d^i$, $x^{R(i)} = d^{R(i)}$.

4 Repeat steps '2' and '3' in an iterative procedure until convergence is achieved.

5 Calculate the rates $m^{R(i)i}$, $bm^{R(i)i}$, $m^{iR(i)}$, b^i, $md^{iR(i)}$ from equations (19), (21), (25), (26), and (27) respectively, and the equivalent terms for $R(i)$. If desired, s^i can be calculated from equation (26), and $s^{R(i)}$ from an equivalent equation. The terms D^i, B^i, and S^i can also be obtained from equations (4), (5), and (3) respectively, so that the figure 2 accounts are now complete. $w^i(t+T)$ can be calculated from equation (28).

6 In order to make a projection for the region i population, the rates d^i, b^i, $bm^{R(i)i}$, $m^{R(i)i}$, and $m^{iR(i)}$ in equation (28) must be projected from some time series analysis.

4.4 *Model using corrected at-risk population rates*

For the model described in the preceding two sections, we defined simple demographic rates which could be applied to base populations. This procedure only tells a part of the story. To illustrate the further complications which have to be considered, let us take the birth flows B^i and TB^i. These terms include births to mothers who have migrated into region i during the period t to $t+T$. This is correct, since these births are additions to the population, and they appear in the appropriate cells of figure 2. However, it does mean that when rates are estimated in relation to these flows, $w^i(t)$ is not the correct base population. What is needed is some estimate of 'population at risk' which would obviously include the inmigrating mothers mentioned above.

Our task now, then, is to estimate corrected at-risk population rates, and to assess resulting changes in the model. The rates may be calculated in the following way: estimate the correct population, from which the events or flows recorded in the numerator can be said to have originated, and then divide by this to obtain the corrected rate.

We shall find that, for births, the at-risk population and the associated rate will have to be incorporated in the model equation itself. For deaths the correct population total appears in the model, but we argue that a better estimate of the rate can be obtained from total deaths and an associated at-risk population. In the case of migration, again the corrected term appears in the model, and usually the basic data is in such a form that no calculation of at-risk population is needed. However, for certain cases, notably where international migration flows are involved, such calculations are necessary, but we shall postpone our discussion of these to a later paper. In section 4.5 below we calculate the at-risk population for births, and note the corresponding modifications to the model. In section 4.6 we do the same thing for deaths. In section 4.7 we summarise the calculation procedure in the revised model.

4.5 *An at-risk population for births, a corrected birth rate, and associated changes in the model*

In dealing with the problem outlined in the previous section, we also find that we resolve the minor inconsistency about the divisor for B^i in our earlier discussions.

It seemed that we should divide by TB^i instead of $w^i(t)$, and it is now convenient to define

$$\beta^i = \frac{B^i}{TB^i} = \frac{TB^i - BD^i - BM^{iR(i)} - BMD^{iR(i)}}{TB^i} . \tag{30}$$

The main problem now becomes one of estimating TB^i, especially when the model is used for projection. We define a birth rate tb^i, calculated from this total figure, as

$$tb^i = \frac{TB^i}{\hat{w}^{iB}} , \tag{31}$$

where $\hat{w}^{iB}$ is the at-risk population for total births. Note that this is a population which refers to the *period* t to $t+T$, not simply the base year t.

This means that the term $b^i w^i(t)$ in equation (28) should be replaced by $\beta^i tb^i \hat{w}^{iB}$. We can recall that the term $BM^{R(i)i}$ should, strictly, have been divided by $TB^{R(i)}$ instead of $w^{R(i)}(t)$, and we can apply a similar argument. We now replace equation (21) by

$$\beta m^{R(i)i} = \frac{BM^{R(i)i}}{TB^{R(i)}} = \frac{BM^{R(i)i}}{tb^{R(i)}\hat{w}^{R(i)B}} , \tag{32}$$

so that the $bm^{R(i)i}w^{R(i)}(t)$ term in equation (28) is replaced by $\beta m^{R(i)i}tb^{R(i)}\hat{w}^{R(i)B}$. Thus the modified model is

$$\begin{aligned} w^i(t+T) = w^i(t) - d^i w^i(t) + \beta^i tb^i \hat{w}^{iB} + \beta m^{R(i)i} tb^{R(i)} \hat{w}^{R(i)B} + m^{R(i)i} w^{R(i)}(t) \\ - m^{iR(i)} w^i(t) - md^{iR(i)} w^i(t) . \end{aligned} \tag{33}$$

Thus it now remains to show how $\hat{w}^{iB}$ (and $\hat{w}^{R(i)B}$) can be calculated. To illustrate the principles involved in defining an at-risk population, we introduce a new type of figure—figure 4. This helps to answer the question of who can be a mother[(10)] giving birth in region i. Vertical lines on the diagram depict time cross-sections $t, t+T$. Horizontal lines divide the region i from the rest of the world $R(i)$. The spatial dimension of the figure is merely a convenient topological transformation of a map[(11)]. Represented on the figure are individual lifelines that move horizontally through time. Vertical shifts represent migrations. For example, lifeline (7) shows someone who survives the period $(t, t+T)$; lifeline (9) shows someone who dies; (1) shows a person who is born in region i and survives there, and (2) a birth followed by a migration from region i to $R(i)$. In the accounts we are concerned with populations at the cross-sections (first row and first column for region i in figure 2), and the recording of demographic events which take place in the period t to $t+T$.[(12)] The 'known' data, shown as shaded cells on figure 2, represent either cross-section populations or events counts on such diagrams.

There are six kinds of mothers we are interested in—types (7) to (12) depicted in figure 4a. In the figure the birth symbols are shown on the lifeline of the mother, and refer to events experienced by the mother. We assume that each type of mother spends a different portion of the total period living in region i and exposed to the risk of giving birth as listed in table 1.

We assume that surviving migrants spend about half of the period in region i and half in region $R(i)$, that persons who die in region i (D^i) are exposed for about half

(10) We use the term 'mother' rather than 'person' or 'population' in the paragraphs that follow, although the model is not explicitly disaggregated by sex. It sounds more natural to talk about mothers in this context, though without fathers there would be no mothers at risk!

(11) See Rees and Wilson (1972a) for an exposition of the three dimensional time-space diagram.

(12) Note. The definitions of the migration events are constrained. A migration out of region i and back again in $(t, t+T)$, lifeline (7), for example, would not be recorded.

the time, and persons who migrate and die spend a quarter of the period in region i and a quarter in region $R(i)$. Mothers who survive in region $R(i)$ ($S^{R(i)}$) or die there ($D^{R(i)}$) are ignored, though some of them may have spent some time in region i (figure 4c). We assume this time is cancelled out by the time spent in the rest of the world by region i survivors (S^i) and persons who die in i (D^i). We also assume that babies born in the period do not reach reproductive age in the same period, that is, that T is less than 15 years (figure 4b).

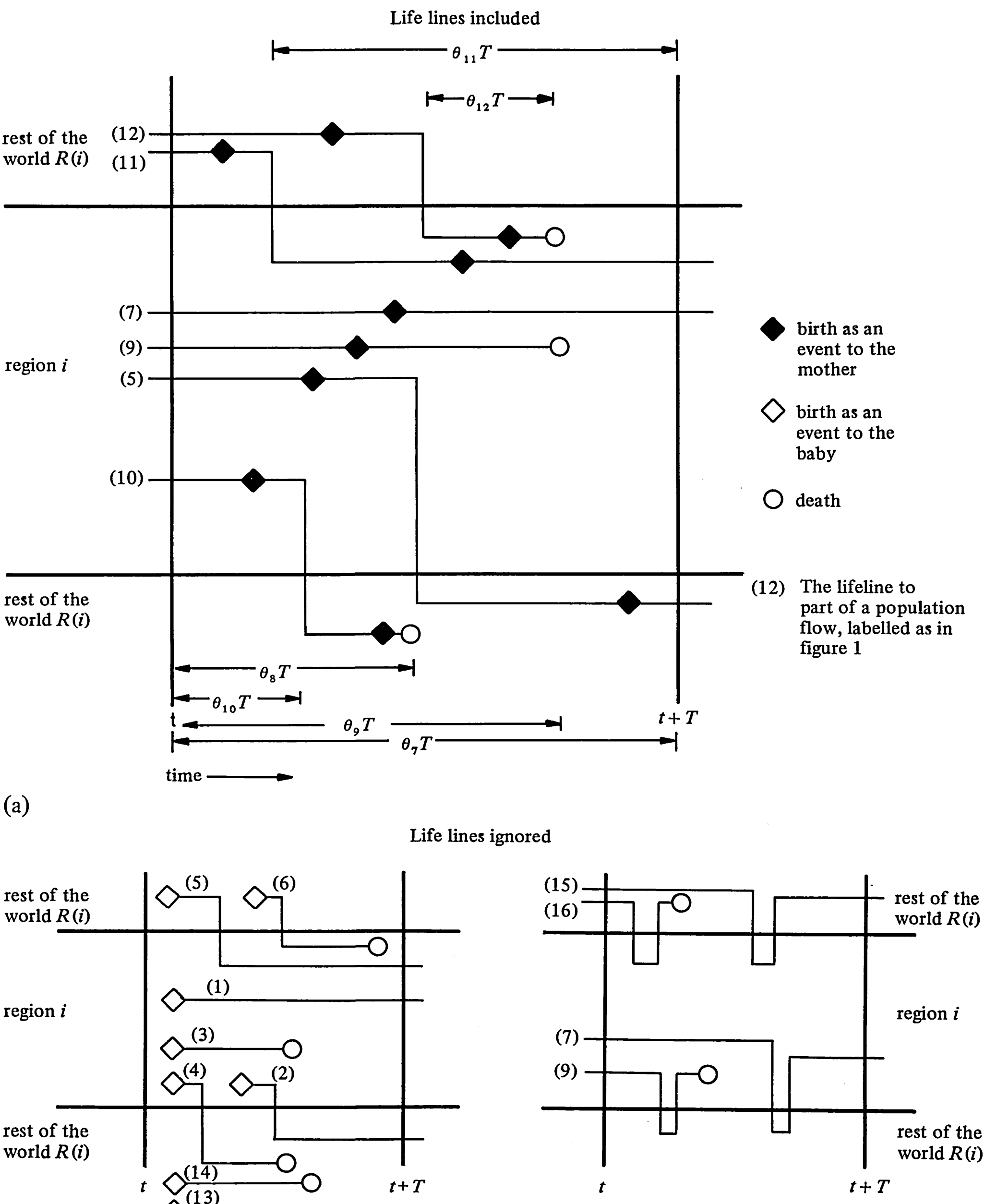

Figure 4. Lifelines involved in population at risk of giving birth. (a) Mothers; (b) babies; (c) multiple movers.

Suppose we now introduce the concept of rates as events per unit time exposed (taking t to $t+T$ as *one* unit). We couple this concept with an assumption that rate-behaviour, or probabilities of events, is the same for all people in a region *for the time they are there.* This concept can be applied to the problem of calculating the population of mothers at risk.

The at-risk population is made up of mothers of types (1) to (6), weighted by the length of exposure of the population flow in region i. The total person-years of exposure to the risk of birth in region i are equal to

$$(\theta_7^i T)S^i + (\theta_9^i T)D^i + (\theta_8^i T)M^{iR(i)} + (\theta_{10}^i T)MD^{iR(i)} + (\theta_{11}^i T)M^{R(i)i} + (\theta_{12}^i T)MD^{R(i)i} .$$

On dividing through by T, the at-risk population in region i is given by

$$\hat{w}^{iB} = \theta_7^i S^i + \theta_9^i D^i + \theta_8^i M^{iR(i)} + \theta_{10}^i MD^{iR(i)} + \theta_{11}^i M^{R(i)i} + \theta_{12}^i MD^{R(i)i} , \tag{34}$$

or, if we substitute for S^i from equation (3),*

$$\hat{w}^{iB} = w^i(t) - (1-\theta_9^i)D^i - (1-\theta_8^i)M^{iR(i)} - (1-\theta_{10}^i)MD^{iR(i)} + (1-\theta_{11}^i)M^{R(i)i} + (1-\theta_{12}^i)MD^{R(i)i} . \tag{35}$$

By adopting the approximations for the θ^i's suggested in table 1, the population at risk of birth in region i becomes

$$\hat{w}^{iB} = w^i(t) - (\tfrac{1}{2})D^i - (\tfrac{1}{2})M^{iR(i)} - (\tfrac{3}{4})MD^{iR(i)} + (\tfrac{1}{2})M^{R(i)i} + (\tfrac{1}{4})MD^{R(i)i} . \tag{36}$$

The total birth rate is therefore

$$tb^i = \frac{TB^i}{w^i(t) - (\frac{1}{2})D^i - (\frac{1}{2})M^{iR(i)} - (\frac{3}{4})MD^{iR(i)} + (\frac{1}{2})M^{R(i)i} + (\frac{1}{4})MD^{R(i)i}} . \tag{37}$$

Table 1.† Terms in an 'at risk' population for births.

Type [(a)]	Population flow	Time exposed in region i	Approximation for θ
(7)	mothers who survive in region i (S^i)	$\theta_7^i T$	1
(9)	mothers who die in region i (D^i)	$\theta_9^i T$	$\frac{1}{2}$
(8)	mothers who outmigrate to region $R(i)$ ($M^{iR(i)}$)	$\theta_8^i T$	$\frac{1}{2}$
(10)	mothers who outmigrate to region $R(i)$ and die there ($MD^{iR(i)}$)	$\theta_{10}^i T$	$(\frac{1}{2})(\frac{1}{2}) = (\frac{1}{4})$
(11)	mothers who inmigrate to region i ($M^{iR(i)}$)	$\theta_{11}^i T$	$\frac{1}{2}$
(12)	mothers who inmigrate to region i and die there ($MD^{R(i)i}$)	$\theta_{12}^i T$	$(\frac{1}{2})(\frac{1}{2}) = (\frac{1}{4})$
(15)	mothers who survive in region $R(i)$ ($S^{R(i)}$)	$\theta_{15}^i T$	0
(16)	mothers who die in region $R(i)$ ($D^{R(i)}$)	$\theta_{16}^i T$	0

(a) The numbering corresponds to that used in figure 1.

† *Note added in proof.* Further investigation suggests that the approximate value for θ of $\frac{1}{4}$ associated with flow (10) should be $\frac{1}{2}$. Our own argument for this was supported by a formal derivation given by E. K. Gilje (personal communication).

* *Note added in proof.* It may be possible to improve this calculation of $\hat{w}^{iB}$ (and the procedure for $\hat{w}^{iD}$ below) by also substituting for D^i from equation (4).

We now have tb^i and $\hat{w}^{iB}$ (and by a similar calculation $tb^{R(i)}$ and $\hat{w}^{R(i)B}$), which may be substituted in the revised model equation (33). When a projection is being made, tb^i has to be projected exogenously, and $\hat{w}^{iB}$ can be calculated from an equation such as (36).

This is all we need to do for computational purposes, but it is of interest, in the interpretation of the effect of introducing at-risk populations for births into the model, to substitute from equation (36) (and the equivalent for $\hat{w}^{R(i)B}$) into the model equation (33). We then get

$$\begin{aligned} w^i(t+T) = {} & w^i(t) - d^i w^i(t) - m^{iR(i)} w^i(t) - md^{iR(i)} w^i(t) \\ & + \beta^i tb^i [w^i(t) - (\tfrac{1}{2})D^i - (\tfrac{1}{2})M^{iR(i)} - (\tfrac{3}{4})MD^{iR(i)} + (\tfrac{1}{2})M^{R(i)i} + (\tfrac{1}{4})MD^{R(i)i}] \\ & + m^{R(i)i} w^{R(i)}(t) + \beta m^{R(i)i} tb^{R(i)} [w^{R(i)}(t) - (\tfrac{1}{2})D^{R(i)} - (\tfrac{1}{2})M^{R(i)i} \\ & - (\tfrac{3}{4})MD^{R(i)i} + (\tfrac{1}{2})M^{iR(i)} + (\tfrac{1}{4})MD^{iR(i)}] . \end{aligned} \tag{38}$$

The terms within the brackets can be defined in terms of rates multiplied by the appropriate rate populations:

$$\begin{aligned} w^i(t+T) = {} & w^i(t) - d^i w^i(t) - m^{iR(i)} w^i(t) - md^{iR(i)} w^i(t) \\ & + \beta^i tb^i [w^i(t) - (\tfrac{1}{2})d^i w^i(t) - (\tfrac{1}{2})m^{iR(i)} w^i(t) - (\tfrac{3}{4})md^{iR(i)} w^i(t) \\ & + (\tfrac{1}{2})m^{R(i)i} w^{R(i)}(t) + (\tfrac{1}{4})md^{R(i)i} w^{R(i)}] + m^{R(i)i} w^{R(i)}(t) \\ & + \beta m^{R(i)i} tb^{R(i)} [w^{R(i)}(t) - (\tfrac{1}{2})d^{R(i)} w^{R(i)}(t) - (\tfrac{1}{2})m^{R(i)i} w^{R(i)}(t) \\ & - (\tfrac{3}{4})md^{R(i)i} w^{R(i)}(t) + (\tfrac{1}{2})m^{iR(i)} w^i(t) + (\tfrac{1}{4})md^{iR(i)} w^i(t)] . \end{aligned} \tag{39}$$

This can be expanded and rearranged to read

$$\begin{aligned} w^i(t+T) = {} & w^i(t) - d^i w^i(t) - m^{iR(i)} w^i(t) - md^{iR(i)} w^i(t) + m^{R(i)i} w^{R(i)}(t) \\ & + \beta^i tb^i w^i(t) - \beta^i tb^i (\tfrac{1}{2}) d^i w^i(t) - \beta^i tb^i (\tfrac{1}{2}) m^{iR(i)} w^i(t) - \beta^i tb^i (\tfrac{3}{4}) md^{iR(i)} w^i(t) \\ & + \beta^i tb^i (\tfrac{1}{2}) m^{R(i)i} w^{R(i)}(t) + \beta^i tb^i (\tfrac{1}{4}) md^{R(i)i} w^{R(i)}(t) + \beta m^{R(i)i} tb^{R(i)} w^{R(i)}(t) \\ & - \beta m^{R(i)i} tb^{R(i)} (\tfrac{1}{2}) d^{R(i)} w^{R(i)}(t) - \beta m^{R(i)i} tb^{R(i)} (\tfrac{1}{2}) m^{R(i)i} w^{R(i)}(t) \\ & - \beta m^{R(i)i} tb^{R(i)} (\tfrac{3}{4}) md^{R(i)i} w^{R(i)}(t) + \beta m^{R(i)i} tb^{R(i)} (\tfrac{1}{2}) m^{iR(i)} w^i(t) \\ & + \beta m^{R(i)i} tb^{R(i)} (\tfrac{3}{4}) md^{iR(i)} w^i(t) . \end{aligned} \tag{40}$$

In the first five terms the simple and corrected rate estimating principles coincide; the next twelve terms show how the population at risk concept can be related back to the populations of the regions at the beginning of the period. What model equation (40) does is to trace the parental origin and place of birth of the persons who are added to the population over the time period, allowing for deaths and migration, and making the assumption that regional or interregional rates for birth, survival, and migration apply, no matter which group of parents is concerned, and that they experience the average birth rate of the region they are located in, when they are located there.

4.6 *An at-risk population for deaths, and an associated death rate*

The way in which a total birth rate can be defined in terms of total births and the population at risk has been described in the previous section. We now describe how a total death rate can be constructed using the population-at-risk concept.

In fact exactly the same argument can be used as was used for the birth rate, except that this time all sixteen population flows must be included. These flows, the average time exposed in the region, and the approximation assumed are listed in table 2.

We assume that people who are born are around for half the period if they survive; if they migrate and survive they spend on average a quarter of the period in i and a quarter in $R(i)$. Babies who are born and die are around for an average of only a quarter of the period, and if they migrate as well they spend only an eighth of the time in each region. Hence the population at risk of death, using the table 2 approximations, is

$$\begin{aligned}\hat{w}^{iD} &= S^i + \tfrac{1}{2}D^i + \tfrac{1}{2}M^{iR(i)} + \tfrac{1}{4}MD^{iR(i)} + \tfrac{1}{4}M^{R(i)i} + \tfrac{1}{4}MD^{R(i)i} + \tfrac{1}{2}B^i + \tfrac{1}{4}BD^i \\ &\quad + \tfrac{1}{4}BM^{iR(i)} + \tfrac{1}{8}BMD^{iR(i)} + \tfrac{1}{4}BM^{R(i)i} + \tfrac{1}{8}BMD^{R(i)i} . \end{aligned} \tag{41}$$

Substituting for S^i we obtain

$$\begin{aligned}\hat{w}^{iD} &= w^i(t) - \tfrac{1}{2}D^i - \tfrac{1}{2}M^{iR(i)} - \tfrac{3}{4}MD^{iR(i)} + \tfrac{1}{2}M^{R(i)i} + \tfrac{1}{4}MD^{R(i)i} + \tfrac{1}{2}B^i + \tfrac{1}{4}BD^i \\ &\quad + \tfrac{1}{4}BM^{iR(i)} + \tfrac{1}{8}BMD^{iR(i)} + \tfrac{1}{4}BM^{R(i)i} + \tfrac{1}{8}BMD^{R(i)i} , \end{aligned} \tag{42}$$

and a similar equation for $\hat{w}^{R(i)D}$. The total death rates are then given by

$$td^i = \frac{TD^i}{\hat{w}^{iD}} , \tag{43}$$

and equivalently for $td^{R(i)}$.

The rates td^i and $td^{R(i)}$ are not the same as d^i and $d^{R(i)}$. d^i is the death rate which applies to $w^i(t)$ (in relation to the first row of figure 2) and this allows for migrants. td^i is worked out in relation to all deaths and an at-risk population. This suggests that d^i is less stable than td^i, and that we should use td^i for projection purposes. Further, since the minor flow events involving death rates in section 3 are very much 'all deaths' events, not just d^i-events, then x^i should be set to td^i if the latter is available.

d^i then becomes redundant in the model. We can see this as follows: given TD^i, and the section 3 procedure for estimating minor flows from td^i, D^i can always be estimated directly using equation (4). To make this plain we can substitute for BD^i, $MD^{R(i)i}$, and $BMD^{R(i)i}$ in equation (4) from the appropriate equation in section 3, (15), (10), and (12):

$$D^i = TD^i - td^i(TB^i - BM^{iR(i)} - BMD^{iR(i)}) - \frac{td^i}{1-td^i}M^{R(i)i} - \frac{td^i}{1-td^i}BM^{R(i)i} , \tag{44}$$

Table 2. Terms in an 'at risk' population for deaths.

Type[(a)]	Population flow	Time exposed in region i	Approximation for θ
(7)	S^i	$\theta_7 T$	1
(9)	D^i	$\theta_9 T$	$\frac{1}{2}$
(8)	$M^{iR(i)}$	$\theta_8 T$	$\frac{1}{2}$
(10)	$MD^{iR(i)}$	$\theta_{10} T$	$\frac{1}{4}$
(11)	$M^{R(i)i}$	$\theta_{11} T$	$\frac{1}{2}$
(12)	$MD^{R(i)i}$	$\theta_{12} T$	$\frac{1}{4}$
(15)	$S^{R(i)}$	$\theta_{15} T$	0
(16)	$D^{R(i)}$	$\theta_{16} T$	0
(1)	B^i	$\theta_1 T$	$\frac{1}{2}$
(3)	BD^i	$\theta_3 T$	$\frac{1}{4}$
(2)	$BM^{iR(i)}$	$\theta_2 T$	$\frac{1}{4}$
(4)	$BMD^{iR(i)}$	$\theta_4 T$	$\frac{1}{8}$
(5)	$BM^{R(i)i}$	$\theta_5 T$	$\frac{1}{4}$
(6)	$BMD^{R(i)i}$	$\theta_6 T$	$\frac{1}{8}$
(13)	$B^{R(i)}$	$\theta_{13} T$	0
(14)	$BD^{R(i)}$	$\theta_{14} T$	0

[(a)] The numbering corresponds to that used in figure 1.

or, substituting for $BMD^{iR(i)}$, from equation (11)

$$D^i = TD^i - td^i\left(TB^i - BM^{iR(i)} - \frac{td^{R(i)}}{1-td^{R(i)}}BM^{iR(i)}\right) - \frac{td^i}{1-td^i}M^{R(i)i} - \frac{td^i}{1-td^i}BM^{R(i)i} \, . \quad (45)$$

Equation (33) can then be written

$$w^i(t+T) = w^i(t) - D^i + \beta^i tb^i \hat{w}^{iB} + bm^{R(i)i} tb^{R(i)} \hat{w}^{R(i)B} + m^{R(i)i} w^{R(i)}(t)$$
$$- m^{iR(i)} w^i(t) - md^{iR(i)} w^i(t) \, , \quad (46)$$

where, now, $md^{iR(i)}$ should be obtained from equation (9) with $x^{R(i)} = td^{R(i)}$ to give, instead of equation (26),

$$md^{iR(i)} = \frac{td^{R(i)}}{1-td^{R(i)}} \, . \quad (47)$$

4.7 *Summary of the model building procedure using corrected birth and death rates*
The procedure resulting from the analysis of sections 4.4–4.6 is a relatively complicated one. It may be useful to collect together the equations given there and to produce a concise statement of the model for region i. We also note that it is easy to write down equivalent equations for a system of N regions, and this is done in Appendix 1.

4.7.1 *The accounts building procedures.*
Step 1. Estimate the following from census and registration records: $w^i(t)$, $w^{R(i)}(t)$, $M^{iR(i)}$, $M^{R(i)i}$, TD^i, $TD^{R(i)}$, $BM^{iR(i)}$, $BM^{R(i)i}$, TB^i, $TB^{R(i)}$.

Step 2. Calculate the at-risk populations for births and for deaths:

$$\hat{w}^{iB} = w^i(t) - \tfrac{1}{2}D^i - \tfrac{1}{2}M^{iR(i)} - \tfrac{3}{4}MD^{iR(i)} + \tfrac{1}{2}M^{R(i)i} + \tfrac{1}{4}MD^{R(i)i} \, , \quad (48)$$

and

$$\hat{w}^{iD} = w^i(t) - \tfrac{1}{2}D^i - \tfrac{1}{2}M^{iR(i)} - \tfrac{3}{4}MD^{iR(i)} + \tfrac{1}{2}M^{R(i)i} + \tfrac{1}{4}MD^{R(i)i} - \tfrac{1}{2}B^i - \tfrac{1}{4}BD^i$$
$$+ \tfrac{1}{4}BM^{iR(i)} + \tfrac{1}{8}BMD^{iR(i)} + \tfrac{1}{4}BM^{R(i)i} + \tfrac{1}{8}BMD^{iR(i)} \, . \quad (49)$$

Initially any unestimated terms could be set to zero, or the population at risk could be set equal to the start-of-period population:

$$\hat{w}^{iB} = w^i(t) - \tfrac{1}{2}M^{iR(i)} + \tfrac{1}{2}M^{R(i)i} \, , \quad (50)$$

or

$$\hat{w}^{iB} = w^i(t) \, ,$$

and

$$\hat{w}^{iD} = w^i(t) - \tfrac{1}{2}M^{iR(i)} + \tfrac{1}{2}M^{R(i)i} - \tfrac{1}{4}BM^{iR(i)} + \tfrac{1}{4}BM^{R(i)i} \, , \quad (51)$$

or

$$\hat{w}^{iD} = w^i(t) \, . \quad (52)$$

Equivalent equations are used to define $\hat{w}^{R(i)B}$ and $\hat{w}^{R(i)D}$. The population at risk of giving birth is not used directly in the account building procedure in the historical context, but is used in calculating birth rates in the model version of the account building algorithm. These rates will be projected and used in the population projection.

Step 3. Calculate td^i from

$$td^i = \frac{TD^i}{\hat{w}^{iD}} , \tag{53}$$

and $td^{R(i)}$ from

$$td^{R(i)} = \frac{TD^{R(i)}}{\hat{w}^{R(i)D}} . \tag{54}$$

Step 4. Calculate $MD^{iR(i)}$, $MD^{R(i)i}$, $BMD^{iR(i)}$, $BMD^{R(i)i}$, and BD^i from the following equations:

$$MD^{iR(i)} = \frac{td^{R(i)}}{1-td^{R(i)}} M^{iR(i)} , \tag{55}$$

$$MD^{R(i)i} = \frac{td^i}{1-td^i} M^{R(i)i} , \tag{56}$$

$$BMD^{iR(i)} = \frac{td^{R(i)}}{1-td^{R(i)}} BM^{iR(i)} , \tag{57}$$

$$BMD^{R(i)i} = \frac{td^i}{1-td^i} BM^{R(i)i} , \tag{58}$$

$$BD^i = td^i(TB^i - BM^{iR(i)} - BMD^{iR(i)}) , \tag{59}$$

with equivalent equations for the corresponding $R(i)$ terms.

Step 5. Once the smaller unknown flows have been calculated the larger unknown flows can be estimated:

$$D^i = TD^i - BD^i - MD^{R(i)i} - BMD^{R(i)i} , \tag{60}$$

$$S^i = w^i(t) - D^i - M^{iR(i)} - MD^{iR(i)} , \tag{61}$$

and

$$B^i = TB^i - BD^i - BM^{iR(i)} - BMD^{iR(i)} . \tag{62}$$

The terms $D^{R(i)}$, $S^{R(i)}$, and $B^{R(i)}$ can be calculated from the region $R(i)$ equivalents of equations (60), (61), and (62).

Step 6. End of period population may now be calculated using

$$w^i(t+T) = S^i + B^i + M^{R(i)i} + BM^{R(i)i} , \tag{63}$$

and the population accounts table has now been completely filled in for the first time.

Steps 2–6 can now be repeated using the first estimate of the population accounts table. They are repeated until the estimate of end of period population converges on a stable value. A best estimate of $w^i(t+T)$ has now been obtained.

4.7.2 *The account building procedure in model form.* In order to project the population of a region forward, it is necessary to express the account building procedure in model form.

Firstly, a set of *rates* must be defined which will yield, when multiplied by the appropriate base populations, projections of the known terms given in *step 1* of the account building procedure. These rates are defined in terms of the population

accounts table and can be measured historically. For population projection purposes, of course, they must themselves be projected into the future[13].
Migration rates are given by

$$m^{iR(i)} = \frac{M^{iR(i)}}{w^i(t)} , \tag{64}$$

$$m^{R(i)i} = \frac{M^{R(i)i}}{w^{R(i)}(t)} . \tag{65}$$

Total death rates have been defined in equations (53) and (54).
Total birth rates are obtained from

$$tb^i = \frac{TB^i}{\hat{w}^{iB}} , \tag{66}$$

$$tb^{R(i)} = \frac{TB^{R(i)}}{\hat{w}^{R(i)B}} . \tag{67}$$

Birth and migration rates are defined as follows as

$$\beta m^{R(i)i} = \frac{BM^{R(i)i}}{TB^{R(i)}} , \tag{68}$$

$$\beta m^{iR(i)} = \frac{BM^{iR(i)}}{TB^i} , \tag{69}$$

though in some cases it may be more convenient to use the simple rates

$$bm^{R(i)i} = \frac{BM^{R(i)i}}{w^{R(i)}(t)} , \tag{70}$$

$$bm^{iR(i)} = \frac{BM^{iR(i)}}{w^i(t)} . \tag{71}$$

The known flows listed at *step 1* can be redefined in terms of the rates and appropriate base populations:

$$M^{iR(i)} = m^{iR(i)}w^i(t) , \tag{72}$$

$$M^{R(i)i} = m^{R(i)i}w^{R(i)}(t) , \tag{73}$$

$$TD^i = td^i\hat{w}^{iD} , \tag{74}$$

$$TD^{R(i)} = td^{R(i)}\hat{w}^{R(i)D} , \tag{75}$$

$$TB^i = tb^i\hat{w}^{iB} , \tag{76}$$

$$TB^{R(i)} = tb^{R(i)}\hat{w}^{R(i)B} , \tag{77}$$

$$\beta M^{iR(i)} = \beta m^{iR(i)}tb^i\hat{w}^{iB} , \tag{78}$$

or

$$BM^{iR(i)} = bm^{iR(i)}w^i(t) , \tag{79}$$

$$BM^{R(i)i} = \beta m^{R(i)i}tb^{R(i)}\hat{w}^{R(i)B} , \tag{80}$$

or

$$BM^{R(i)i} = bm^{R(i)i}w^{R(i)}(t) , \tag{81}$$

[13] Methods and models for doing this are discussed in later papers (Rees and Wilson, 1972a, and Wilson and Rees, 1972b).

We can express the constituent terms of the $w^i(t+T)$ equation [equation (63)] in model form as follows:

$$S^i = w^i(t) - \left[td^i\hat{w}^{iD} - td^i\left(tb^i\hat{w}^{iB} - \beta m^{iR(i)}tb^i\hat{w}^{iB} - \frac{td^{R(i)}}{1-td^{R(i)}}\beta m^{iR(i)}tb^i\hat{w}^{iB}\right)\right.$$
$$\left. - \frac{td^i}{1-td^i}m^{R(i)i}w^{R(i)}(t) - \frac{td^i}{1-td^i}\beta m^{R(i)i}tb^{R(i)}\hat{w}^{R(i)B}\right] - m^{iR(i)}w^i(t)$$
$$- \frac{td^{R(i)}}{1-td^{R(i)}}m^{iR(i)}w^i(t) , \tag{82}$$

$$B^i = tb^i\hat{w}^{iB} - td^i\left(tb^i\hat{w}^{iB} - \beta m^{iR(i)}tb^i\hat{w}^{iB} - \frac{td^{R(i)}}{1-td^{R(i)}}\beta m^{iR(i)}tb^i\hat{w}^{iB}\right) - \beta m^{iR(i)}tb^i\hat{w}^{iB}$$
$$- \frac{td^{R(i)}}{1-td^{R(i)}}\beta m^{iR(i)}tb^i\hat{w}^{iB} , \tag{83}$$

$$M^{R(i)i} = m^{R(i)i}w^{R(i)}(t) , \tag{84}$$

and

$$BM^{R(i)i} = \beta m^{R(i)i}tb^{R(i)}\hat{w}^{R(i)B} . \tag{85}$$

Equations (82)–(85) can be combined to give a fully explicit model equation for $w^i(t+T)$. This equation differs from previous model equations [for example, equation (40)] by including only rates calculated from known flows.

Equation (77) says that all that is necessary in population projection is to be able to say how the rates $m^{iR(i)}$, $m^{R(i)i}$, td^i, $td^{R(i)}$, tb^i, $tb^{R(i)}$, $\beta m^{iR(i)}$, and $\beta m^{R(i)i}$ will change in the future, and to be able to generate iteratively, via the accounts building procedure, estimates of the populations at risk.

4.7.3 *The projection procedure.* The projection calculations proceed a little differently from those of the account building in the historical case and are sequenced as follows:

Step 1. Projections should be made of the 'known flow' rates defined in the preceding section (4.7.2), and $w^i(t)$ and $w^{R(i)}(t)$ should be set equal to the population of the base year from which the projections will start.

Step 2. Calculate the migration flow terms using equations (72) and (73). If the simple rate versions of the birth and migration rates are used, then the birth and migration flows may be calculated as well using equations (81) and (80). Otherwise the calculation of birth and migration terms is postponed to a later step.

Step 3. Calculate the at-risk populations as in *step 2* of the accounting building procedures, setting the unestimated terms to zero. Initially equation (50) will be used for population at risk of birth; at subsequent iterations equation (48) will be used. For population at risk of dying initially, the following equation will be used

$$\hat{w}^{iD} = w^i(t) - \tfrac{1}{2}M^{iR(i)} + \tfrac{1}{2}M^{R(i)i} , \tag{86}$$

and subsequently equation (49), unless the birth and migration terms have already been estimated. In that case, equation (51) will be used initially and (49) in subsequent iterations.

Step 4. Calculate TD^i and $TD^{R(i)}$ from equations (74) and (75), TB^i and $TB^{R(i)}$ from (76) and (77), and $BM^{R(i)i}$ and $BM^{iR(i)}$ from (78) and (80), unless the birth and migration terms have already been calculated.

Step 5. Calculate the small unknown flows $MD^{iR(i)}$, $MD^{R(i)i}$, $BMD^{iR(i)}$, $BMD^{R(i)i}$, and BD^i, and the equivalent $R(i)$ terms as in *step 4* of the accounts building procedure.

Step 6. Calculate the larger unknown flows D^i, S^i, and B^i and their $R(i)$ equivalents as in *step 5* of the accounts building procedure.

Step 7. The projected population can be calculated using equation (64), and the projected population accounts table is now complete for the first time.

Step 8. Cycle through *steps 3–7* until convergence is achieved and best estimates of the projected populations of region i and $R(i)$ are obtained.

4.8 *Comment*
It is not immediately clear how to compare the models presented here with those developed by such workers as Rogers (1966, 1968) using a matrix format. Such relationships are explored in a forthcoming paper which translates the results presented here into matrix terms (Rees and Wilson, 1972b).

5 A worked example
In order to demonstrate that the account building procedures described in preceding sections are empirically feasible, and the accounts so constructed can form the basis of population models, we describe a simple example of a set of population accounts for the West Riding of Yorkshire, UK for the census period 1961 to 1966. We emphasise that this is to illustrate concepts only, and later, the same methods will be applied to an age–sex disaggregated case using the methods described in Wilson and Rees (1972a).

Table 3. The population accounts table for the West Riding of Yorkshire, 1961–1966 census period: conceptual terms.

Situation at time t	Category at time t	Situation at time $t+T$						Totals population terms
		Yorkshire, West Riding region i		Rest of England and Wales region $N(i)$		Rest of the world region R		
		Category at time $t+T$						
		alive in 1966	died 1961–1966	alive in 1966	died 1961–1966	alive in 1966	died 1961–1966	resident [enumerated]
Yorkshire, West Riding region i	alive in 1961	S^i	D^i	$M^{iN(i)}$	$MD^{iN(i)}$	M^{iR}	MD^{iR}	$w^{i2}(t)$ $[w^{i1}(t)]$
	born 1961–1966	B^i	BD^i	$BM^{iN(i)}$	$BMD^{iN(i)}$	BM^{iR}	BMD^{iR}	TB^i
Rest of England and Wales region $N(i)$	alive in 1961	$M^{N(i)i}$	$MD^{N(i)i}$	$S^{N(i)}$	$D^{N(i)}$	$M^{N(i)R}$	$MD^{N(i)R}$	$w^{N(i)2}(t)$ $[w^{N(i)1}(t)]$
	born 1961–1966	$BM^{N(i)i}$	$BMD^{N(i)i}$	$B^{N(i)}$	$BD^{N(i)}$	$BM^{N(i)R}$	$BMD^{N(i)R}$	$TB^{N(i)}$
Rest of the world region R	alive in 1961	$M^{R(i)}$	$MD^{R(i)}$	$M^{RN(i)}$	$MD^{RN(i)}$	S^R	D^R	$w^{R2}(t)$ $[w^{R1}(t)]$
	born 1961–1966	$BM^{R(i)}$	$BMD^{R(i)}$	$BM^{RN(i)}$	$BMD^{RN(i)}$	B^R	BD^R	TB^R
Population terms (a)								
Accounts total		w^{i4}		$w^{N(i)4}$		w^{R4}		
Total inflated resident		w^{i3}	TD^i	$w^{N(i)3}$	$TD^{N(i)}$	w^{R3}	TD^R	
(Resident)		(w^{i2})		$(w^{N(i)2})$		(w^{R2})		
[Enumerated]		$[w^{i1}]$		$[w^{N(i)1}]$		$[w^{R1}]$		

(a) The column population totals refer to time $t+T$.

Regions

i Yorkshire, West Riding (1961 definition).

$N(i)$ Rest of England and Wales, that is, England and Wales minus Yorkshire, West Riding.

R Rest of the world, that is, world minus England and Wales. (We might have labelled this $W(N)$ to be consistent with the $N(i)$ notation, but we retain R to be consistent with earlier $R(i)$ notation.)

Table 3 sets out the population accounts table in conceptual form and table 4 contains the estimated population flows that correspond with the conceptual accounts. Initially, only those flows which can be estimated from existing sources of demographic data are entered in table 4. All these known flows involve use of some estimation procedure. For example, no figures are given directly for persons born outside the West Riding who inmigrate and survive there in 1966. These have to be estimated from data on the number of children in inmigrant families. All the migration figures have to be inflated by a correction factor to allow for the known underenumeration of the Sample Census of 1966. Outmigration totals to the rest of the world involve the use of several different estimation techniques for the different constituent zones (Scotland, Northern Ireland, Isle of Man and Channel Islands, Irish Republic, rest of the world outside the British Isles). Some of these methods are described in Rees (1971), and a forthcoming paper (Rees, 1972b) will describe in detail how one goes about the difficult task of matching theoretical information requirements with the information available from published sources. Here we assume that the numbers in table 4 are the best estimates we are likely to get for these flows.

The problem now is to fill in the empty cells with estimates for these flows, using the methods described in the preceding sections of the paper. A preliminary step is to decide what is unreasonable to attempt to estimate. In table 4, any attempt to estimate the population of the rest of the world or the population flows within that region are obviously infeasible. When building historical accounts like those for 1961-1966 this means that a different assumption about the death rate experienced by outmigrants has to be made than that posited earlier. To estimate the MD^{iR} term we use the West Riding death rate, arguing that this would reflect more accurately the experience of that flow of persons than the death rate for the vast region that is the rest of the world. The death rate experienced by persons migrating from a small

Table 4. The population accounts table for the West Riding of Yorkshire, 1961-1966, showing known flows. (Source of data: General Register Office, 1968.)

Situation at time t	Category at time t	Situation at time $t+T$: Yorkshire, West Riding region i — Category at time $t+T$: alive in 1966	died 1961-1966	Rest of England and Wales region $N(i)$: alive in 1966	died 1961-1966	Rest of the world region R: alive in 1966	died 1961-1966	Totals population terms: resident [enumerated]
Yorkshire, West Riding region i	alive in 1961			168207		55328		3650586 [3644582]
	born 1961-1966			7391		4221		333321
Rest of England and Wales region $N(i)$	alive in 1961	137183				831390		42453962 [42459966]
	born 1961-1966	7327				57031		3925236
Rest of the world region R	alive in 1961	58804		1012460				
	born 1961-1966	2509		44358				
Totals: Accounts total								
Inflated resident		3741153		44069974				
(Resident)		(3688720)	226529	(43452330)	2537801			
[Enumerated]		[3686500]		[43449010]				

zone into a vastly larger one is likely to be more closely approximated by the smaller zone's rate than the larger's, especially if we know that the migration is very selective with respect to destinations within the larger zone.

The calculation now proceeds in a number of steps which roughly parallel those listed in section 4.7 for the corrected rate model building procedure. Initially, $\hat{w}^{iD}$ and $\hat{w}^{iB}$ are taken as $w^i(t)$. The results are presented for each iteration in turn so that the effects of iterating can be seen. Of course, in this example, there are three regions which we have designated i, $N(i)$, and R. The reader will easily be able to make the appropriate adjustments from the $[i, R(i)]$ system described in earlier sections.

Death rates, first iteration

Step 1. The death rates are calculated for each zone in the system for which a set of full accounts is to be generated. In our case these zones are region i, the West Riding, and region $N(i)$, the rest of England and Wales, but not region R, the rest of the world. The death rates are calculated by dividing total deaths in the region by the opening stock population

$$td^k = \frac{TD^k}{w^k(t)} \; ; \tag{87}$$

for $k = i$

$$td^i = \frac{226\,529}{3\,650\,586} = 0{\cdot}062\,053 = \frac{62{\cdot}053}{1000} \; , \tag{88}$$

and for $k = N(i)$

$$td^{N(i)} = \frac{2\,537\,801}{42\,453\,962} = 0{\cdot}059\,778 = \frac{59{\cdot}778}{1000} \; . \tag{89}$$

Migration and death factors, first iteration

Step 2. The death rates are used to calculate factors by which the migration totals are multiplied to yield estimates of the migration and death flows. We calculate

$$\frac{td^k}{1-td^k} \; ; \tag{90}$$

for $k = i$

$$\frac{td^i}{1-td^i} = 0{\cdot}066\,158 = \frac{66{\cdot}158}{1000} \; , \tag{91}$$

and

$$\frac{td^{N(i)}}{1-td^{N(i)}} = 0{\cdot}063\,578 = \frac{63{\cdot}578}{1000} \; . \tag{92}$$

These factors are then applied as follows:

$$MD^{iN(i)} = \frac{td^{N(i)}}{1-td^{N(i)}} M^{iN(i)} = 0{\cdot}63\,578 \times 168\,207 = 10\,694 \; , \tag{93}$$

$$MD^{iR} = \frac{td^i}{1-td^i} M^{iR} = 0{\cdot}066\,158 \times 55\,328 = 3660 \; , \tag{94}$$

$$BMD^{iN(i)} = \frac{td^{N(i)}}{1-td^{N(i)}} BM^{iN(i)} = 0{\cdot}63\,578 \times 7391 = 470 \; , \tag{95}$$

and so on. Terms M^{iR}, BM^{iR}, $M^{N(i)i}$, $BM^{N(i)i}$, $M^{R(i)}$, and $BM^{R(i)}$ are multiplied by $td^i/(1-td^i)$ to yield MD^{iR}, BMD^{iR}, $MD^{N(i)i}$, $BMD^{N(i)i}$, $MD^{R(i)}$, and $BMD^{R(i)}$. Terms

$M^{iN(i)}$, $BM^{iN(i)}$, $M^{N(i)R}$, $BM^{N(i)R}$, $M^{RN(i)}$, and $BM^{RN(i)}$ are multiplied by $td^{N(i)}/(1-td^{N(i)})$ to yield the equivalent migration and death terms. Table 5 shows the state of the accounts at the end of this step. As an estimate of the infant death rate the total death used here is, of course, likely to be inaccurate. However, this inaccuracy is not crucial in this case as we are concerned only with accounting for and projecting the population in aggregate.

Birth and death terms, first iteration

Step 3. The next step in constructing the accounts is to calculate the birth and death terms for any fully explicit region. For any region k,

$$BD^k = td^k\left(TB^k - \sum_{j\neq k} BM^{kj} - \sum_{j\neq k} BMD^{kj}\right) , \tag{96}$$

where the summation includes all regions, $j \neq k$.

For $k = i$, equation (10) becomes

$$\begin{aligned} BD^i &= td^i(TB^i - BM^{iN(i)} - BMD^{iN(i)} - BM^{iR} - BMD^{iR}) \\ &= (0\cdot072053)(333321 - 7391 - 470 - 4221 - 279) \\ &= (0\cdot062653)(320960) = 19916 . \end{aligned} \tag{97}$$

For $k = N(i)$, equation (84) becomes

$$\begin{aligned} BD^{N(i)} &= TD^{N(i)}(TB^{N(i)} - BM^{N(i)i} - BMD^{N(i)i} - BM^{N(i)R} - BMD^{N(i)R}) \\ &= 0\cdot059778(3925236 - 7327 - 485 - 57031 - 3626) \\ &= 0\cdot059778(3856767) = 230547 . \end{aligned} \tag{98}$$

Table 5. The population accounts table for the West Riding of Yorkshire, 1961-1966, at the end of *step 2*. (Source of data: General Register Office, 1968.)

Situation at time t	Category at time t	Situation at time $t+T$: Yorkshire, West Riding region i — Category at time $t+T$: alive in 1966	died 1961-1966	Rest of England and Wales region $N(i)$: alive in 1966	died 1961-1966	Rest of the world region R: alive in 1966	died 1961-1966	Totals population terms: resident [enumerated]
Yorkshire, West Riding region i	alive in 1961			168207	10694	55328	3660	3650586
	born 1961-1966			7391	470	4221	279	333321
Rest of England and Wales region $N(i)$	alive in 1961	137183	9076			831390	52858	42453962
	born 1961-1966	7327	485			57031	3626	3925236
Rest of the world region R	alive in 1961	58804	3890	1012460	63370			
	born 1961-1966	2509	166	44358	2820			
Totals:								
Accounts total								
Inflated resident								
(Resident)			226529		2537801			
[Enumerated]								

Deaths in situ terms, first iteration
Step 4. Having calculated the birth and death term, we are now able to calculate the deaths *in situ* term. In general this is given by the following relation,

$$D^k = TD^k - BD^k - \sum_{j \neq k} MD^{kj} - \sum_{j \neq k} BMD^{kj} \, . \tag{99}$$

For $k = i$, equation (13) becomes

$$\begin{aligned} D^i &= TD^i - BD^i - MD^{N(i)i} - BMD^{N(i)i} - MD^{R(i)} - BMD^{R(i)} \\ &= 226529 - 19916 - 9076 - 485 - 3890 - 166 = 192996 \, . \end{aligned} \tag{100}$$

For the rest of England and Wales, region $N(i)$, the deaths term is given by

$$\begin{aligned} D^{N(i)} &= TD^{N(i)} - BD^{N(i)} - MD^{iN(i)} - BMD^{iN(i)} - MD^{RN(i)} - BMD^{RN(i)} \\ &= 2537801 - 230547 - 10694 - 570 - 64370 - 2820 = 2228900 \, . \end{aligned} \tag{101}$$

The survival terms, first iteration
Step 5. The final step in building the accounting table is to work out the survivors term and the surviving births term. For survivors, the following equation is used

$$S^k = w^k(t) - D^k - \sum_{j \neq k} M^{kj} - \sum_{j \neq k} MD^{kj} \, . \tag{102}$$

For the West Riding, region i, this translates as

$$\begin{aligned} S^i &= w^i(t) - D^i - M^{iN(i)} - MD^{iN(i)} - M^{iR} - MD^{iR} \\ &= 3650586 - 192996 - 168207 - 10794 - 55328 - 3660 = 3219701 \, . \end{aligned} \tag{103}$$

For region $N(i)$ this becomes

$$\begin{aligned} S^{N(i)} &= w^{N(i)}(t) - D^{N(i)} - M^{N(i)i} - MD^{N(i)i} - M^{N(i)R} - MD^{N(i)R} \\ &= 42453962 - 2228900 - 137183 - 9076 - 831390 - 52850 = 39194555 \, . \end{aligned} \tag{104}$$

Surviving infants born in the period are calculated using

$$B^k = TB^k - BD^k - \sum_{j \neq k} BM^{kj} - \sum_{j \neq k} BMD^{kj} \, , \tag{105}$$

which for West Yorkshire is

$$\begin{aligned} B^i &= TB^i - BD^i - BM^{iN(i)} - BMD^{iN(i)} - BM^{iR} - BMD^{iR} \\ &= 333321 - 19916 - 7391 - 470 - 4221 - 279 = 301044 \, , \end{aligned} \tag{106}$$

and for the rest of England and Wales is

$$\begin{aligned} B^{N(i)} &= TB^{N(i)} - BD^{N(i)} - BM^{N(i)i} - BMD^{N(i)i} - BM^{N(i)R} - BMD^{N(i)R} \\ &= 3925236 - 230547 - 7327 - 485 - 57031 - 3626 = 3626220 \, . \end{aligned} \tag{107}$$

Calculation of the closing stock populations, first iteration
Step 6. The final step in the sequence is the calculation of the new end-of-period populations using the general equation

$$w^k(t+T) = S^k + B^k + \sum_{j \neq k} M^{jk} + \sum_{j \neq k} BM^{jk} \, . \tag{108}$$

In the West Riding case this is

$$\begin{aligned} w^i(t+T) &= S^i + B^i + M^{N(i)i} + BM^{N(i)i} + M^{R(i)} + BM^{R(i)} \\ &= 3219701 + 301044 + 137183 + 7327 + 58804 + 2509 = 3726568 \, , \end{aligned} \tag{109}$$

and for the rest of England and Wales we have the following version of equation (97)

$$\begin{aligned} w^{N(i)}(t+T) &= S^{N(i)}+B^{N(i)}+M^{iN(i)}+BM^{iN(i)}+M^{RN(i)}+BM^{RN(i)} \\ &= 39\,194\,555+2\,626\,220+168\,207+7391+1\,002\,460+44\,358 \\ &= 44\,053\,191\,. \end{aligned} \tag{110}$$

Table 6 shows the state of the accounts after the first six steps.

Calculation of populations at risk of dying, second iteration
Step 7. Conceptually these accounts can be improved by applying death rates calculated using population at risk of dying as the base population. The population at risk of dying, or $\hat{w}^{iD}$, is calculated by adding up all the flow terms, weighted by the average proportion of the period that they are exposed in the region in question, in the population accounting matrix (sections 4.4–4.6). Table 7 sets out the calculation for our three-region system for the two populations at risk we need to calculate.

The population at risk of dying in the period lies between the time t population and the first estimate of the time $t+T$ population as might be expected, but are significantly more than the mid-point populations which are used in conventional demographic analysis as the base populations in rate calculations (Bogue, 1969, p.119). The use of mid-point populations is justified on exactly the same basis as we have justified our population-at-risk definitions. The discrepancy still requires elucidation.

Table 6. The population accounts table for the West Riding of Yorkshire, 1961–1966, at the end of *step 6.* (Source of data: General Register Office, 1968.)

Situation at time t	Category at time t	Situation at time $t+T$: Yorkshire, West Riding region i — Category at time $t+T$: alive in 1966	died 1961–1966	Rest of England and Wales region $N(i)$: alive in 1966	died 1961–1966	Rest of the world region R: alive in 1966	died 1961–1966	Totals population terms: resident [enumerated]
Yorkshire, West Riding region i	alive in 1961	3219701	192996	168207	10694	55328	3660	3650586
	born 1961–1966	301044	19916	7391	470	4221	279	333321
Rest of England and Wales region $N(i)$	alive in 1961	137183	9076	39194555	2228900	831390	52858	42453962
	born 1961–1966	7327	485	3626220	230547	57031	3626	3925236
Rest of the world region R	alive in 1961	58804	3890	1012460	63370			
	born 1961–1966	2509	166	44358	2820			
Totals: Accounts total		3726568		44053191				
Inflated resident (Resident) [Enumerated]			226529		2537801			

Death rates, second iteration

Step 8. New death rates are calculated using $\hat{w}^{iD}$ and $\hat{w}^{N(i)D}$ as the base populations:

$$td^i = \frac{TD^i}{\hat{w}^{iD}} = \frac{226\,529}{3\,693\,828} = 0{\cdot}061\,325 = \frac{61{\cdot}325}{1000}\,, \tag{111}$$

and

$$td^{N(i)} = \frac{TD^{N(i)}}{\hat{w}^{N(i)D}} = \frac{2\,537\,801}{43\,318\,573{\cdot}125} = 0{\cdot}058\,585 = \frac{58{\cdot}585}{1000}\,. \tag{112}$$

These death rates are slightly lower than the ones produced initially in *step 1.*

Table 7. Calculation of the populations at risk of dying, 1961–1966. (Source of data: General Register Office, 1968.)

Term	Flow	Fraction for region i	Constituent of $\hat{w}^{iD}$	Fraction for region $N(i)$	Constituent of $\hat{w}^{N(i)D}$
col.1	col.2	col.3	col.4 = col.3 x col.2	col.5	col.6 = col.2 x col.5
S^i	3219701	1	3219701	0	0
D^i	192996	$\frac{1}{2}$	96498	0	0
$M^{iN(i)}$	168207	$\frac{1}{2}$	84103·5	$\frac{1}{2}$	84103·5
$MD^{iN(i)}$	10694	$\frac{1}{4}$	2673·5	$\frac{1}{4}$	2673·5
M^{iR}	55328	$\frac{1}{2}$	27664	0	0
MD^{iR}	3660	$\frac{1}{4}$	915	0	0
B^i	301044	$\frac{1}{2}$	150522	0	0
BD^i	19916	$\frac{1}{4}$	4979	0	0
$BM^{iN(i)}$	7391	$\frac{1}{4}$	1847·75	$\frac{1}{4}$	1847·75
$BMD^{iN(i)}$	470	$\frac{1}{8}$	58·75	$\frac{1}{8}$	58·75
BM^{iR}	4221	$\frac{1}{4}$	1055·25	0	0
BMD^{iR}	279	$\frac{1}{8}$	34·875	0	0
$M^{N(i)i}$	137183	$\frac{1}{2}$	68591·5	$\frac{1}{2}$	68591·5
$MD^{N(i)i}$	9076	$\frac{1}{4}$	2269	$\frac{1}{4}$	2269
$S^{N(i)}$	39194555	0	0	1	39194555
$D^{N(i)}$	2228900	0	0	$\frac{1}{2}$	1114450
$M^{N(i)R}$	831390	0	0	$\frac{1}{2}$	415695
$MD^{N(i)R}$	52858	0	0	$\frac{1}{4}$	13214·5
$BM^{N(i)i}$	7327	$\frac{1}{4}$	1831·75	$\frac{1}{4}$	1831·75
$BMD^{N(i)i}$	485	$\frac{1}{8}$	60·625	$\frac{1}{8}$	60·625
$B^{N(i)}$	3626220	0	0	$\frac{1}{2}$	1813110
$BD^{N(i)}$	230547	0	0	$\frac{1}{4}$	57636·75
$BM^{N(i)R}$	57031	0	0	$\frac{1}{4}$	14257·75
$BMD^{N(i)R}$	3626	0	0	$\frac{1}{8}$	453·25
$M^{R(i)}$	58804	$\frac{1}{2}$	29402	0	0
$MD^{R(i)}$	3890	$\frac{1}{4}$	972·5	0	0
$M^{RN(i)}$	1012460	0	0	$\frac{1}{2}$	506230
$MD^{RN(i)}$	63370	0	0	$\frac{1}{4}$	16092·5
S^R		0	0	0	0
D^R		0	0	0	0
$BM^{R(i)}$	2509	$\frac{1}{4}$	627·25	0	0
$BMD^{R(i)}$	166	$\frac{1}{8}$	20·75	0	0
$BM^{RN(i)}$	44358	0	0	$\frac{1}{4}$	11089·5
$BMD^{RN(i)}$	2820	0	0	$\frac{1}{8}$	352·5
B^R		0	0	0	0
BD^R		0	0	0	0
Totals			3693828		43318573·125

Steps 9–13, second iteration.

These steps repeat the calculations of *steps 2–6*, using the new death rates. The results of this second iteration are given in table 8 (the upper figures in the cells of flows that change in the iteration). New populations at risk can then be calculated, new death rates computed, and a new population accounting matrix produced. If table 8 is compared with the results of the first iteration (table 5) some small changes can be seen to have occurred. Each of the terms that depend on the death rate directly is lower in table 8 (the *MD, BMD,* and *BD* terms). The D^i and $D^{N(i)}$ terms are consequently larger and *in situ* survivors, S^i and $S^{N(i)}$, smaller in number. On the other hand, the number of surviving infants is larger. The resulting change in the estimated population at the end of the period is only 114 (increase) out of 3726528 in the case of the West Riding and only 343 (decrease) out of 44053191 in the case of the rest of England and Wales.

The third iteration produces virtually no changes in either flow numbers or population totals and the iteration sequence is halted there.

Table 8. The population accounts table for the West Riding of Yorkshire, 1961–1966, at the end of the second iteration and at the end of the third iteration.

Situation at time t	Category at time t	Situation at time $t+T$						Totals
		Yorkshire, West Riding region i		Rest of England and Wales region $N(i)$		Rest of the world region R		
		Category at time $t+T$						
		alive in 1966	died 1961–1966	alive in 1966	died 1961–1966	alive in 1966	died 1961–1966	
Yorkshire,	alive in	3219570	193398	168207	10468	55328	3615	3650586
West Riding	1961	3219511	193397		10468		3615	
region i	born	301289	19684	7391	460	4221	276	333321
	1961–	301289	19684		460		276	
	1966							
Rest of	alive in	137183	8962	39189434	2235255	831390	51738	42453962
England and	1961		8963	39189542	2235144		51740	
Wales	born	7327	479	3630998	225852	57031	3549	3925236
region $N(i)$	1961–		479	3630890	225960		3549	
	1966							
Rest of	alive in	58804	3842	1012460	63006			
the world	1961		3842		63008			
region R	born	2509	164	44358	2760			
	1961–		164		2761			
	1966							
Totals:								
Accounts total		3726682		44052848				
		3726683		44052848				
Inflated resident		3741153						
(Resident)		3668720	226529		2537801			
[Enumerated]		3686500						

Note. The upper figure in any cell represents the state of the population accounts table at the end of the second iteration, and the lower figure the state at the end of the third iteration. Where only one figure is given this refers to a flow that does not change through the iteration procedure.

References

Bogue, D. J., 1969, *The Principles of Demography* (John Wiley, New York).

General Register Office, 1968, *Sample Census 1966. England and Wales. Migration Regional Report. Yorkshire and Humberside Region* (HMSO, London).

Rees, P. H., 1971, "How to make the Rogers' version of the demographic model operational for an English or Welsh county in about one hundred easy steps", Working Paper Number 8, Department of Geography, University of Leeds.

Rees, P. H., 1972a, "The distribution of social groups within cities: models and accounts", in *London Papers in Regional Science, Volume 3. Patterns and Processes in Urban and Regional Systems,* Ed. A. G. Wilson (Pion, London), pp.165–216.

Rees, P. H., 1972b, "Data and related estimation problems for spatial demographic analysis in the UK", Working Paper, Department of Geography, University of Leeds (forthcoming).
Rees, P. H., Wilson, A. G., 1972a, "The time and space framework of population accounting and projection", Working Paper, Department of Geography, University of Leeds (forthcoming).
Rees, P. H., Wilson, A. G., 1972b, "A matrix formulation of accounts and models for spatial demographic analysis", Working Paper, Department of Geography, University of Leeds (forthcoming).
Rogers, A., 1966, "Matrix methods of population analysis", *Journal of the American Institute of Planners,* 32, 40-44.
Rogers, A., 1968, *Matrix Analysis of Inter-regional Population Growth and Distribution* (University of California Press, Berkeley).
Stone, R., 1971, *Demographic Accounting and Model Building* , OECD Technical Reports in Education and Development, Paris.
Wilson, A. G., 1972, "Multiregional models of population structure and some implications for a dynamic-residential location model", in *London Papers in Regional Science, Volume 3. Patterns and Processes in Urban and Regional Systems,* Ed. A. G. Wilson (Pion, London), pp.217-242.
Wilson, A. G., Rees, P. H., 1972a, "Accounts and models for spatial demographic analysis II: Age/sex disaggregated populations", Working Paper, Department of Geography, University of Leeds (forthcoming).
Wilson, A. G., Rees, P. H., 1972b, "Models of regional birth and death rates", Working Paper, Department of Geography, University of Leeds (forthcoming).

APPENDIX I

The equation system for N regions

The equations in the main body of the paper have generally been expressed for a two region system: a region i and a rest of the world region, region $R(i)$. We list the main equations of section 4.7 for region i in a regional system consisting of regions $1, 2, ..., i, ..., j, ..., N$:

Populations at risk

$$\hat{w}^{iB} = w^i(t) - \tfrac{1}{2}D^i - \tfrac{1}{2}\sum_{j \neq i} M^{ij} - \tfrac{3}{4}\sum_{j \neq i} MD^{ij} + \tfrac{1}{2}\sum_{j \neq i} M^{ji} + \tfrac{1}{4}\sum_{j \neq i} MD^{ji} , \tag{A1}$$

$$\begin{aligned}\hat{w}^{iD} = {} & w^i(t) - \tfrac{1}{2}D^i - \tfrac{1}{2}\sum_{j \neq i} M^{ij} - \tfrac{3}{4}\sum_{j \neq i} MD^{ij} + \tfrac{1}{2}\sum_{j \neq i} M^{ji} + \tfrac{1}{4}\sum_{j \neq i} MD^{ji} - \tfrac{1}{2}B^i - \tfrac{1}{4}BD^i \\ & + \tfrac{1}{4}\sum_{j \neq i} BM^{ij} + \tfrac{1}{8}\sum_{j \neq i} BMD^{ij} + \tfrac{1}{4}\sum_{j \neq i} BM^{ji} + \tfrac{1}{8}\sum_{j \neq i} BMD^{ij} .\end{aligned} \tag{A2}$$

Total death rate

$$td^i = \frac{TD^i}{\hat{w}^{iD}} \text{ as before.} \tag{A3}$$

Migrants who die

$$MD^{ij} = \frac{td^j}{1 - td^i} M^{ij} . \tag{A4}$$

Migrating infants who die

$$BMD^{ij} = \frac{td^j}{1 - td^j} BM^{ij} . \tag{A5}$$

Infants who die in situ

$$BD^i = td^i \left(TB^i - \sum_{j \neq i} BM^{ij} - \sum_{j \neq i} BMD^{ij} \right) . \tag{A6}$$

Persons who die

$$D^i = TD^i - BD^i - \sum_{j \neq i} MD^{ji} - \sum_{j \neq i} BMD^{ji} . \tag{A7}$$

Survivors

$$S^i = w^i(t) - D^i - \sum_{j \neq i} M^{ij} - \sum_{j \neq i} MD^{ij} . \tag{A8}$$

Surviving infants

$$B^i = TB^i - BD^i - \sum_{j \neq i} BM^{ij} - \sum_{j \neq i} BMD^{ij} . \tag{A9}$$

End of period population

$$w^i(t+T) = S^i + B^i + \sum_{j \neq i} M^{ji} + \sum_{j \neq i} BM^{ji} . \tag{A10}$$

Migration rates

$$m^{ij} = \frac{M^{ij}}{w^i(t)} . \tag{A11}$$

Total birth rates

$$tb^i = \frac{TB^i}{\hat{w}^{iB}(t)} \text{ as before.} \tag{A12}$$

Infant migration rates

$$\beta m^{ij} = \frac{BM^{ij}}{TB^{i}} \; . \tag{A13}$$

Model equation for end of period population

$$\begin{aligned}
w^{i}(t+T) = w^{i}(t) - \Bigg[& td^{i}\hat{w}^{iD} - td^{i}\left(tb^{i}\hat{w}^{iB} - \sum_{j \neq i}\beta m^{ij} tb^{i}\hat{w}^{iB} - \sum_{j \neq i}\frac{td^{j}}{1-td^{j}}\beta m^{ij} tb^{i}\hat{w}^{iB}\right) \\
& - \sum_{j \neq i}\frac{td^{i}}{1-td^{i}} m^{ji} w^{i}(t) - \sum_{j \neq 1}\frac{td^{i}}{1-td^{i}}\beta m^{ji} tb^{j}\hat{w}^{jB}\Bigg] - \sum_{j \neq i} m^{ij} w^{i}(t) \\
& - \sum_{j \neq i}\frac{td^{j}}{1-td^{j}} m^{ij} w^{i}(t) + tb^{i}\hat{w}^{iB} - td^{i}\Bigg(tb^{i}\hat{w}^{iB} - \sum_{j \neq i}\beta m^{ij} tb^{i}\hat{w}^{iB} \\
& - \sum_{j \neq i}\frac{td^{j}}{1-td^{j}}\beta m^{ij} tb^{i}\hat{w}^{iB}\Bigg) - \sum_{j \neq i}\beta m^{ij} tb^{i}\hat{w}^{iB} - \sum_{j \neq i}\frac{td^{j}}{1-td^{j}}\beta m^{ij} tb^{i}\hat{w}^{iB} \\
& + \sum_{j \neq i} m^{ji} w^{j}(t) + \sum_{j \neq i}\beta m^{ji} tb^{j}\hat{w}^{jB} \; .
\end{aligned} \tag{A14}$$

Note that the regional system must partition the whole world and that this may lead to difficulties with a last region external to the system of interest. Equations (A1) to (A14) then have to be modified. The modifications are left to the reader in relation to his particular problem.