Working Paper 23

ACCOUNTS AND MODELS FOR SPATIAL DEMOGRAPHIC ANALYSIS I: AGGREGATE POPULATIONS*

P.H. Rees and A.G. Wilson

Paper presented to the I.B.G. Population Study Group Symposium on Population Dynamics in Northern England, University of Durham, 21st-23rd September, 1972. Not to be quoted without permission.

* The authors would like to acknowledge the support of S.S.R.C. for part of this work.

Department of Geography,
University of Leeds.

September, 1972.

Working Paper 23

Accounts and models for spatial demographic analysis I: aggregate populations

P.H. Rees and A.G. Wilson

Contents

Accounts and models for spatial demographic analysis I: aggregate population

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 multi-region 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.

Multi-regional 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 (1972-A). In this paper, we try to draw the two threads together, and our aim is to present a multi-region 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 which 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 we show 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 paper, Wilson and Rees (1972A), we apply the same methods to age and sex disaggregated populations. In section 2 we introduce the accounting concepts, and in section 3 we indicate how missing cells in the accounts tables can be filled, and in section 4 we indicate how models can be developed. 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 to do this in principle 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

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

where

$w(t)$	=	population at time t
$D(t,t+T)$	=	deaths in period t to $t+T$
$B(t,t+T)$	=	births in period to t $t+T$
$I(t,t+T)$	=	in-migration in period t to $t+T$*
$O(t,t+T)$	=	out-migration in period t to $t+T$.

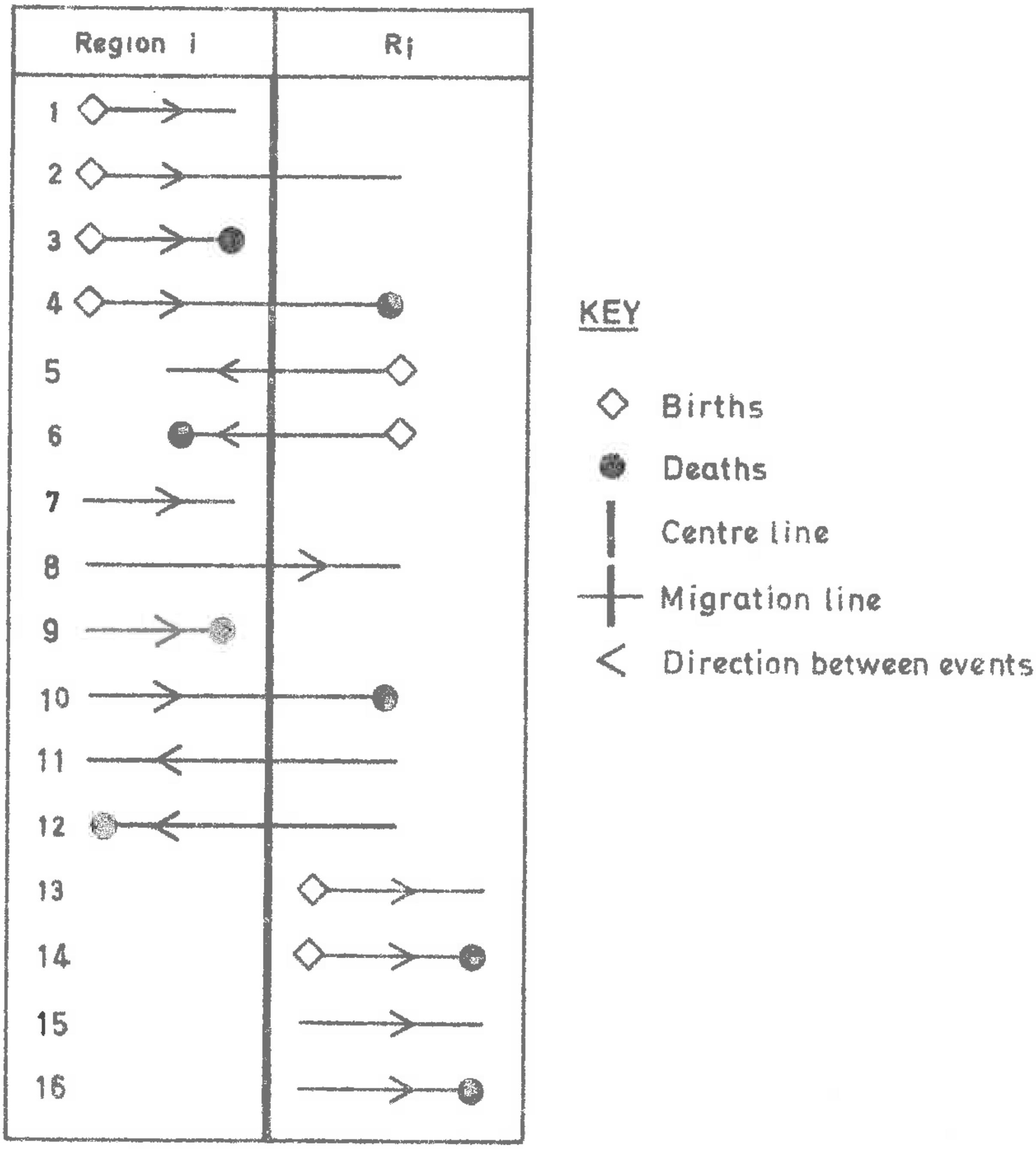
When the time period involved is clear from the content, we shall drop the designation $(t,t+T)$ on flow variables for simplicity.

As noted earlier, our main emphasis is to be the development of a multi-region model. Thus, at the outset, we develop a multi-regional framework. Suppose we have a system of N regions, $i = 1,2, \ldots N$. (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 16 types of demographic event as depicted in Figure 1.

---

* N.B. 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.

Fig. 1


Region i
Rj
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
KEY
Births
Deaths
Centre line
Migration line
Direction between events

The types are as follows (all events in period t to t+T):

(1) Birth in i and survival
(2) Birth in i plus migration to R_i and survival
(3) Birth in i, death in i
(4) Birth in i, migration to R_i, death in R_i
(5) Birth in R_i, migration to i and survival
(6) Birth in R_i, migration to i, death in i
(7) Survival in i
(8) Migration to R_i and survival
(9) Death in i
(10) Migration to R_i and death
(11) Migration to i and survival
(12) Migration to i and death
(13) Birth in R_i and survival
(14) Birth in R_i, death in R_i
(15) Survival in R_i
(16) Death in R_i

To deal with the multi-region case more formally, we need to extend the notation introduced after equation (1). We do this as follows:

$w^i(t)$	=	population in region i at t*
D^i	=	deaths in region i of people in i at the beginning of the period in period t to t+T**
MD^{ij}	=	deaths of people who have migrated from i and die in j
B^i	=	births in i, resident in i at end of period
BD^i	=	births in i, death in i
BM^{ij}	=	births in i, plus migration to j from i and survival in j.
BMD^{ij}	=	birth in i, migration from i to j, death in j
M^{ij}	=	migration from i to j
S^i	=	survival in 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 16 flows of Figure 1 in this notation as follows:

(1) B^i
(2) BM^{iRi}
(3) BD^i
(4) BMD^{iRi}
(5) BM^{Rii}
(6) BMD^{Rii}
(7) S^i
(8) M^{iRi}

---

* This is for any i of course. Also, $w^{Ri}(t)$ would be the population of R_i at t.

** t to t+T is implied for all following flow definitions.

(9) D^i
(10) MD^{iR_i}
(11) $M^{R_i i}$
(12) $MD^{R_i i}$
(13) B^{R_i}
(14) BD^{R_i}
(15) S^{R_i}
(16) D^{R_i}

Following Stone (1971), as developed in Rees (1972-A)* we can incorporate these flows in the accounting framework shown in Figure 2.

Figure 2

Fig. 2

		Region i		Region R_i		
	t+T / t	Alive at t+T (Survivors)	Died in t to t+T	Alive at t+T (Survivors)	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}	

* The accounting table presented here (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 multi-region expansion of Stone's table.

where we have introduced:

TB^i = total observed births in region i in the period

TD^i = total observed deaths in region i in the period

and similarly TB^{R_i} and TD^{R_i} for 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 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 in region i (TB^i) and region R_i (TB^{R_i}) respectively over the period.

The first and third columns of the table tell us about the origins of the populations of region 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:

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

The shaded cells in Figure 2 denote flows which are usually available in historical census data or registration records. 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:

$$D^i = TD^i-BD^i-MD^{R_ii}-BMD^{R_ii} \tag{4}$$

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_ii}, 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 (4) into (3), and from (3) and (5) into (2), to obtain

$$
\begin{aligned}
w^i(t+T) \quad = \quad & w^i(t)-(TD^i-BD^i-MD^{R_i i}-BMD^{R_i i}) \\
& +(TB^i-BD^i-BM^{iR_i}-BMD^{iR_i})+BM^{R_i i} \\
& +M^{R_i i} \\
& -(M^{iR_i}+MD^{iR_i}) \qquad (6)
\end{aligned}
$$

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

New population = Old population
-(deaths in old population)
+(surviving non-migrating births, and surviving in-migrating infants)
+(surviving in-migrants)
-(out-migrants) (7)

Since this accounting relationship takes note of availability of information it is the one we shall usually use. 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 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.

Figure 3

Region at time t		Region at time t+T: 1 Alive	1 Died		i Alive	i Died		j Alive	j Died		N Alive	N Died		Totals
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		

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:

$$M^{iR_i},\ M^{R_i i},\ BM^{iR_i},\ BM^{R_i i},\ TB^i,\ TB^{R_i},\ TD^i,\ \text{and}\ TD^{R_i}.^*$$

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 which is 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 out-migrate 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.**

To proceed, we need death rates for i and R_i. We shall later discuss two ways of estimating these, the best of which 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

---

* The methods for doing this estimation are explained in another paper (Rees, 1972-B).

** 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.

are omitted in the second definition. In the rest of this section, we will 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_ii}, BMD^{iR_i}, BMD^{R_ii}, BD^i, and BD^{R_i}.

Using our general principle, we have

$$MD^{iR_i} = x^{R_i}(M^{iR_i}+MD^{iR_i}) \tag{8}$$

since $M^{iR_i} + MD^{iR_i}$ is the total out-migration 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

$$MD^{R_ii} = \frac{x^i}{1-x^i} M^{R_ii} \tag{10}$$

$$BMD^{iR_i} = \frac{x^{R_i}}{1-x^{R_i}} BM^{iR_i} \tag{11}$$

and

$$BMD^{R_ii} = \frac{x^i}{1-x^i} BM^{R_ii} \tag{12}$$

The term BD^i could actually be cancelled from eqution (6), but in disaggregated versions it re-appears. Also, it makes interpretation of the relevant equations easier. So we retain the term and calculate it as follows. Re-arranging 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, using (13)

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

Similarly,

$$BD^{R_i} \quad = \quad x^{R_i}(TB^{R_i} - BM^{R_i i} - BMD^{R_i i}) \tag{16}$$

B_i can now be calculated from (5) using known flows, BD^i from (15) and $BMD^{R_i i}$ from (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 D^{S_i} similarly.*

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.

So, we now 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

$$BM^{iR_i} \quad = \quad \frac{M^{iR_i}}{w^i(t)} \cdot TB^i \tag{17}$$

with a similar equation for $BM^{R_i i}$.

Usually, of course, we will 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. 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.

---

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

$$D^i \quad = \quad x^i(S^i + D^i)$$

$$\quad = \quad x^i(w^i(t) - M^{iR_i} - MD^{iR_i})$$

using equation (3). So, re-arranging equation (4),

$$BD^i \quad = \quad TD^i - x^i(w^i(t) - M^{iR_i} - MD^{iR_i}) - MD^{R_i i} - BMD^{R_i i}$$

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,

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

as an in-migration 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 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

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

Using equations (18) - (21), equation (2) can then 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, then $w^i(t+2T)$ 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 out-migration rate, and b^i, $bm^{R_i i}$ and $m^{R_i i}$ as in (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

$$md^{iR_i} = \frac{d^{R_i}}{1-d^{R_i}} \quad m^{iR_i} \tag{27}$$

where we have replaced x^{R_i} by the death rate being used here, d^{R_i} (defined in the R_i analogue of eduation (24)). Then, equation (23) can be written

$$\begin{aligned} w^i(t+T) = \; & w^i(t)-d^iw^i(t)+b^iw^i(t) \\ & +bm^{R_ii}w^{R_i}(t)+m^{R_ii}w^{R_i}(t) \\ & -m^{iR_i}w^i(t)-md^{iR_i}w^i(t) \end{aligned} \tag{28}$$

Then, provided we can project the rates d^i, b^i, bm^{R_ii}, m^{R_ii}, m^{iR_i} (and calculate md^{iR_i} from equation (27)), then 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 equation (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_ii}, TD^i, TD^{R_i}, BM^{iR_i}, BM^{R_ii}, 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_ii} and BMD^{R_ii} and the equivalent R_i terms to zero.

(3) Calculate MD^{iR_i}, MD^{R_ii}, BMD^{R_ii}, BMD^{R_ii}, 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) add (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 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 in-migrating 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 can 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. However, 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 <u>An at risk population for births, a corrected birth rate, and associated changes in the model</u>

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 to project. We now 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 <u>period</u> 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

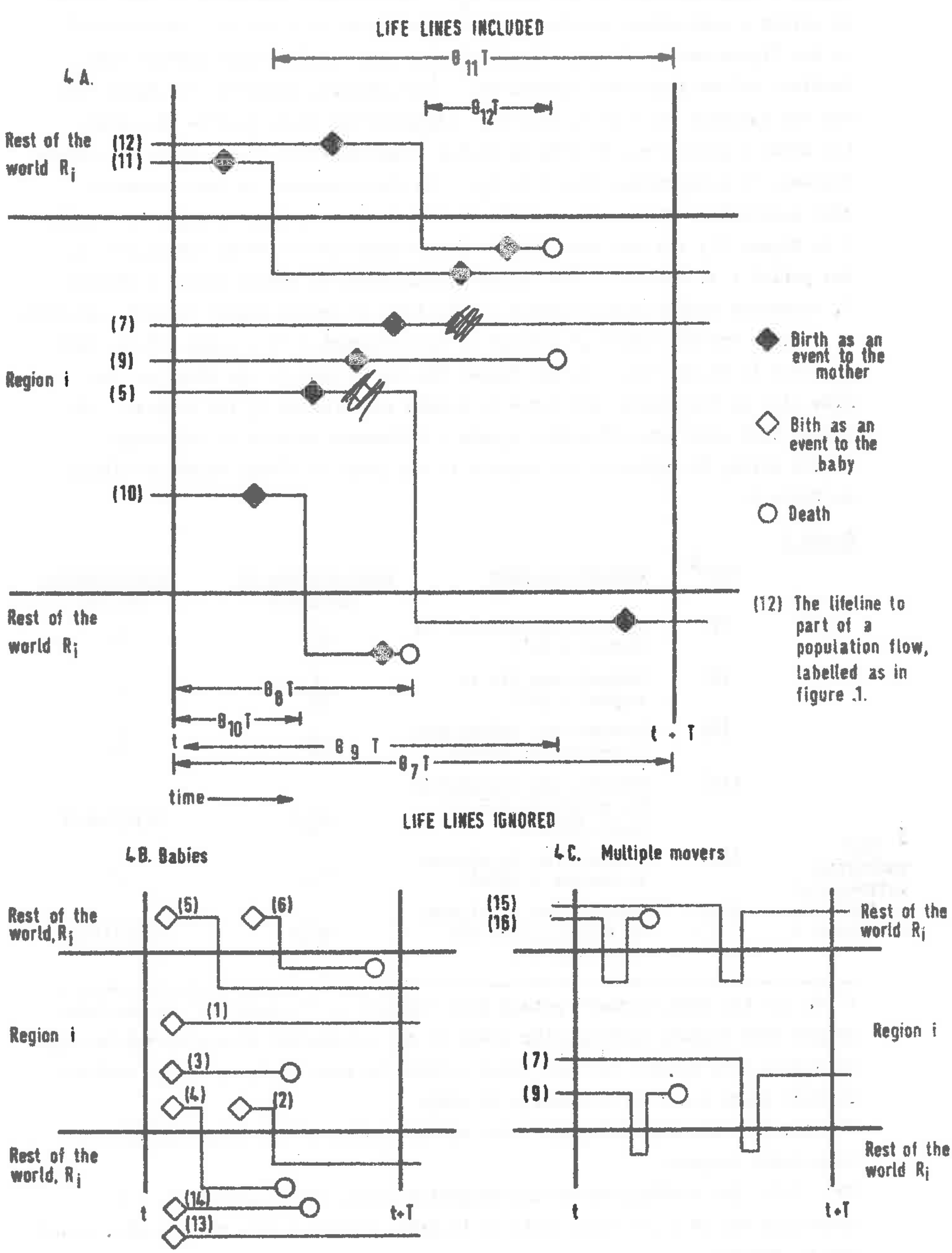
$$\beta m^{R_i i} = \frac{BM^{R_i i}}{TB^{R_i}} = \frac{BM^{R_i i}}{tb^{R_i}\hat{w}^{R_i B}} \tag{32}$$

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

LIFELINES INVOLVED IN POPULATION AT RISK OF GIVING BIRTH Fig. 4



of who can be a mother* 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.** 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 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.*** 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 life line 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.

Table 1

Type**	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 out-migrate to region R_i (M^{iRi})	$\theta_8^i T$	$\frac{1}{2}$
(10)	Mothers who out-migrate to region R_i and die there (MD^{iRi})	$\theta_{10}^i T$	$(\frac{1}{2})(\frac{1}{2})=(\frac{1}{4})$
(11)	Mothers who in-migrate to region i (M^{Rii})	$\theta_{11}^i T$	$\frac{1}{2}$
(12)	Mothers who in-migrate to region i and die there (MD^{Rii})	$\theta_{12}^i T$	$(\frac{1}{2})(\frac{1}{2})=(\frac{1}{4})$

** The numbering corresponds to that used in Figure 1.

* 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!

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

*** N.B. The definitions of the migration events are constrained. A migration out of i and back again is (t,t+T), lifeline (7), for example, would not be recorded.

Table 1 (Cont:)

Type	Population flow	Time exposed in region i	Approximation for θ
(15)	Mothers who survive in region R_i (S^{Ri})	$\theta^i_{15}T$	0
(16)	Mothers who die in region R_i (D^{Ri})	$\theta^i_{16}T$	0

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 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^{Ri}) or die there (D^{Ri}) 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).

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, are 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

$$(\theta^i_7T)S^i+(\theta^i_9T)D^i+(\theta^i_8T)M^{iRi}+(\theta^i_{10}T)MD^{iRi}+(\theta^i_{11}T)M^{Rii}+(\theta^i_{12}T)MD^{Rii}$$

Dividing through by T the at risk population in region i is

$$\hat{w}^{iB} \quad = \quad \theta^i_7S^i+\theta^i_9D^i+\theta^i_8M^{iRi}+\theta^i_{10}MD^{iRi}+\theta^i_{11}M^{Rii}+\theta_{12}MD^{Rii} \qquad (34)$$

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

$$\hat{w}^{iB} \quad = \quad w^i(t)-(1-\theta^i_9)D^i-(1-\theta^i_8)M^{iRi}-(1-\theta^i_{10})MD^{iRi}+(1-\theta^i_{11})M^{Rii}$$
$$+(1-\theta^i_{12})MD^{Rii} \qquad (35)$$

Adopting the approximations for the θ^i's suggested in Table 1 the population at risk of birth in region i becomes

$$\hat{w}^{iB} = w^i(t)-(\tfrac{1}{2})D^i-(\tfrac{1}{2})M^{iRi}-(\tfrac{3}{4})MD^{iRi}+(\tfrac{1}{2})M^{Rii}+(\tfrac{1}{4})MD^{Rii} \quad (36)$$

The total birth rate is therefore

$$tb^i = \frac{TB^i}{w^i(t)-(\tfrac{1}{2})D^i-(\tfrac{1}{2})M^{iRi}-(\tfrac{3}{4})MD^{iRi}+(\tfrac{1}{2})M^{Rii}+(\tfrac{1}{4})MD^{Rii}} \quad (37)$$

We now have tb^i and $\hat{w}^{iB}$ (and by a similar calculation tb^{Ri} and $\hat{w}^{RiB}$) which can 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 affect of introducing at risk populations for births into the model to substitute from equation (36) (and the equivalent for $\hat{w}^{RiB}$) into the model equation (33). We then get

$$\begin{aligned} w^i(t+T) = {} & w^i(t)-d^iw^i(t)-m^{iRi}w^i(t)-md^{iRi}w^i(t) \\ & +\beta^i tb^i[w^i(t)-(\tfrac{1}{2})D^i-(\tfrac{1}{2})M^{iRi}-(\tfrac{3}{4})MD^{iRi}+(\tfrac{1}{2})M^{Rii}+(\tfrac{1}{4})MD^{Rii}] \\ & +m^{Rii}w^{Ri}(t) \\ & +\beta m^{Rii}tb^{Ri}[w^{Ri}(t)-(\tfrac{1}{2})D^{Ri}-(\tfrac{1}{2})M^{Rii}-(\tfrac{3}{4})MD^{Rii} \\ & +(\tfrac{1}{2})M^{iRi}+(\tfrac{1}{4})MD^{iRi}] \end{aligned} \quad (38)$$

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

$$\begin{aligned} w^i(t+T) = {} & w^i(t)-d^iw^i(t)-m^{iRi}w^i(t)-md^{iRi}w^i(t) \\ & +\beta^i tb^i[w^i(t)-(\tfrac{1}{2})d^iw^i(t)-(\tfrac{1}{2})m^{iRi}w^i(t) \\ & -(\tfrac{3}{4})md^{iRi}w^i(t)+(\tfrac{1}{2})m^{Rii}w^{Ri}(t)+(\tfrac{1}{4})md^{Rii}w^{Ri}(t)] \\ & +m^{Rii}w^{Ri}(t) \\ & +\beta m^{Rii}tb^{Ri}[w^{Ri}(t)-(\tfrac{1}{2})d^{Ri}w^{Ri}(t)-(\tfrac{1}{2})m^{Rii}w^{Ri}(t) \\ & -(\tfrac{3}{4})md^{Rii}w^{Ri}(t)+(\tfrac{1}{2})m^{iRi}w^i(t)+(\tfrac{1}{4})md^{iRi}w^i(t)] \end{aligned} \quad (39)$$

This can be expanded and re-arranged to read

$$
\begin{aligned}
w^{i}(t+T) \quad = \quad & w^{i}(t) && -d^{i}w^{i}(t)-m^{iRi}w^{i}(t)-md^{iRi}w^{i}(t) \\
& +m^{Ri\,i}w^{Ri}(t) && \\
& +\beta^{i}tb^{i}w^{i}(t) && -\beta^{i}tb^{i}(\tfrac{1}{2})d^{i}w^{i}(t) \\
& && -\beta^{i}tb^{i}(\tfrac{1}{2})m^{iRi}w^{i}(t) \\
& && -\beta^{i}tb^{i}(\tfrac{3}{4})md^{iRi}w^{i}(t) \\
& +\beta^{i}tb^{i}(\tfrac{1}{2})m^{Ri\,i}w^{Ri}(t) && \\
& +\beta^{i}tb^{i}(\tfrac{1}{4})md^{Ri\,i}w^{Ri}(t) && \\
& +\beta m^{Ri\,i}tb^{Ri}w^{Ri}(t) && -\beta m^{Ri\,i}tb^{Ri}(\tfrac{1}{2})d^{Ri}w^{Ri}(t) \\
& && -\beta m^{Ri\,i}tb^{Ri}(\tfrac{1}{2})m^{Ri\,i}w^{Ri}(t) \\
& && -\beta m^{Ri\,i}tb^{Ri}(\tfrac{3}{4})md^{Ri\,i}w^{Ri}(t) \\
& +\beta m^{Ri\,i}tb^{Ri}(\tfrac{1}{2})m^{iRi}w^{i}(t) && \\
& +\beta m^{Ri\,i}tb^{Ri}(\tfrac{3}{4})md^{iRi}w^{i}(t) && \qquad (40)
\end{aligned}
$$

In the first five terms the simple and corrected rate estimating principles coincide; the next 12 terms show how the population at risk concept can be related back to the populations of the regions at the beginning of the period. The accounting table which corresponds to the birth terms in the model is shown in Figure 5. Each of the terms in the table is represented in the model as follows:

Number of babies born in region i to persons on a particular type of lifeline (or members of a particular population flow) who survive in region i to the end of the period = (Probability of surviving in region i)
x (Birth rate of region i)
x (Probability of parent being present

in region i over time period)
x (Probability of parents being in population flow)
x (Number of persons in relevant region at beginning of period) (41)

or as

Number of babies born in region R_i to persons in a particular population flow who migrate to region i and survive there at the end of the period = (Probability of migrating and surviving from R_i to i)
x (Birth rate of Region R_i)
x (Probability of parent being present in region R_i over time period)
x (Probability of parents being population flow)
x (Numbers of persons in relevant region at beginning of period). (42)

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 what the group of parents concerned, and that they experience the average birth rate of the region they are located in, when they are located there.

4.6 <u>An at risk population for deaths, and an associated death rate</u>

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 16 population flows must be included. These

Figure 5 Parental origin, place of birth and state at t+T of babies born in the period

Original population	Parental population flow	Region i				Region R^i				Region of birth
		Alive in i	Died in i	Alive in R^i	Died in R^i	Alive in R^i	Died in R^i	Alive in i	Died in i	State at t+T of babies
$w^i(t)$	S^i (7)	$B^i(S^i)$	$BD^i(S^i)$	$BM^{iR_i}(S^i)$	$BMD^{iR_i}(S^i)$	$B^{R_i}(S^i)$	$BD^{R_i}(S^i)$	$BM^{R_i i}(S^i)$	$BMD^{R_i i}(S^i)$	$TB(S^i)$
	D^i (9)	$B^i(D^i)$	$BD^i(D^i)$	$BM^{iR_i}(D^i)$	$BMD^{iR_i}(D^i)$	$B^{R_i}(D^i)$	$BD^{R_i}(D^i)$	$BM^{R_i i}(D^i)$	$BMD^{R_i i}(D^i)$	$TB(D^i)$
	M^{iR_i} (8)	$B^i(M^{iR_i})$	$BD^i(M^{iR_i})$	$BM^{iR_i}(M^{iR_i})$	$BMD^{iR_i}(M^{iR_i})$	$B^{R_i}(M^{iR_i})$	$BD^{R_i}(M^{iR_i})$	$BM^{R_i i}(M^{iR_i})$	$BMD^{R_i i}(M^{iR_i})$	$TB(M^{iR_i})$
	MD^{iR_i} (10)	$B^i(MD^{iR_i})$	$BD^i(MD^{iR_i})$	$BM^{iR_i}(MD^{iR_i})$	$BMD^{iR_i}(MD^{iR_i})$	$B^{R_i}(MD^{iR_i})$	$BD^{R_i}(MD^{iR_i})$	$BM^{R_i i}(MD^{iR_i})$	$BMD^{R_i i}(MD^{iR_i})$	$TB(MD^{iR_i})$
$w^{R_i}(t)$	S^{R_i} (15)	$B^i(S^{R_i})$	$BD^i(S^{R_i})$	$BM^{iR_i}(S^{R_i})$	$BMD^{iR_i}(S^{R_i})$	$B^{R_i}(S^{R_i})$	$BD^{R_i}(S^{R_i})$	$BM^{R_i i}(S^{R_i})$	$BMD^{R_i i}(S^{R_i})$	$TB(S^{R_i})$
	D^{R_i} (16)	$B^i(D^{R_i})$	$BD^i(D^{R_i})$	$BM^{iR_i}(D^{R_i})$	$BMD^{iR_i}(D^{R_i})$	$B^{R_i}(D^{R_i})$	$BD^{R_i}(D^{R_i})$	$BM^{R_i i}(D^{R_i})$	$BMD^{R_i i}(D^{R_i})$	$TB(D^{R_i})$
	$M^{R_i i}$ (11)	$B^i(M^{R_i i})$	$BD^i(M^{R_i i})$	$BM^i(M^{R_i i})$	$BMD^{iR_i}(M^{R_i i})$	$B^{R_I}(M^{R_i i})$	$BD^{R_i}(M^{R_i i})$	$BM^{R_i i}(M^{R_i i})$	$BMD^{R_i i}(M^{R_i i})$	$TB(M^{R_i i})$
	$MD^{R_i i}$ (12)	$B^i(MD^{R_i i})$	$BD^i(MD^{R_i i})$	$BM^i(MD^{R_i i})$	$BMD^{iR_i}(MD^{R_i i})$	$B^{R_i}(MD^{R_i i})$	$BD^{R_i}(MD^{R_i i})$	$BM^{R_i i}(MD^{R_i i})$	$BMD^{R_i i}(MD^{R_i i})$	$TB(MD^{R_i i})$
	Birth totals	B^i	BD^i	BM^i	BMD^{iR_i}	B^{R_i}	BD^{R_i}	$BM^{R_i i}$	$BMD^{R_i i}$	
*	Recorded births	TB^i				TB^{R_i}				TB^*

flows, the average time exposed in the region, and the approximation assumed are listed in Table 2.

Table 2

Type*	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^{iRi}	$\theta_8 T$	$\frac{1}{2}$
(10)	MD^{iRi}	$\theta_{10} T$	$\frac{1}{4}$
(11)	M^{Rii}	$\theta_{11} T$	$\frac{1}{2}$
(12)	MD^{Rii}	$\theta_{12} T$	$\frac{1}{4}$
(15)	S^{Ri}	$\theta_{15} T$	0
(16)	D^{Ri}	$\theta_{16} T$	0
(1)	B^{i}	$\theta_1 T$	$\frac{1}{2}$
(3)	BD^{i}	$\theta_3 T$	$\frac{1}{4}$
(2)	BM^{iRi}	$\theta_2 T$	$\frac{1}{4}$
(4)	BMD^{iRi}	$\theta_4 T$	$\frac{1}{8}$
(5)	BM^{Rii}	$\theta_5 T$	$\frac{1}{4}$
(6)	BMD^{Rii}	$\theta_6 T$	$\frac{1}{8}$
(13)	B^{Ri}	$\theta_{13} T$	0
(14)	BD^{Ri}	$\theta_{14} T$	0

* The numbering corresponds to that used in Figure 1.

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

$$\hat{w}^{iD} = S^i+\tfrac{1}{2}D^i+\tfrac{1}{2}M^{iR_i}+\tfrac{1}{4}MD^{iR_i}$$
$$+\tfrac{1}{4}M^{R_ii}+\tfrac{1}{4}MD^{R_ii}+\tfrac{1}{2}B^i+\tfrac{1}{4}BD^i$$
$$+\tfrac{1}{4}BM^{iR_i}+\tfrac{1}{8}BMD^{iR_i}+\tfrac{1}{4}BM^{R_ii}$$
$$+\tfrac{1}{8}BMD^{R_ii} \tag{43}$$

Substituting for S^i we obtain

$$\hat{w}^{iD} = w^i(t)-\tfrac{1}{2}D^i-\tfrac{1}{2}M^{iR_i}-\tfrac{3}{4}MD^{iR_i}$$
$$+\tfrac{1}{2}M^{R_ii}+\tfrac{1}{4}MD^{R_ii}+\tfrac{1}{2}B^i+\tfrac{1}{4}BD^i$$
$$+\tfrac{1}{4}BM^{iR_i}+\tfrac{1}{8}BMD^{iR_i}+\tfrac{1}{4}BM^{R_ii}$$
$$+\tfrac{1}{8}BMD^{R_ii} \tag{44}$$

and a similar equation for $\hat{w}^{R_iD}$. The total death rates are then given by

$$td^i = TD^i/\hat{w}^{iD} \tag{45}$$

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, then D^i can always be estimated directly using equation (4). To make this plain, we can substitute for BD^i, MD^{R_ii} and BMD^{R_ii} in (4) from the appropriate equation in section 3, (15), (10) and (12):

$$D^i = TD^i - td^i(TB^i - BM^{iRi} - BMD^{iRi}) - \frac{td^i}{1-td^i} M^{Rii} - \frac{td^i}{1-td^i} BM^{Ri^i} \quad (46)$$

or, substituting for BMD^{iRi}, from (11)

$$D^i = TD^i - td^i(TB^i - BM^{iRi} - \frac{td^{Ri}}{1-td^{Ri}} BM^{iRi}) - \frac{td^i M^{Rii}}{1-td^i} - \frac{td^i}{1-td^i} BM^{Rii} \quad (47)$$

Equation (33) can then be written

$$\begin{aligned} w^i(t+T) = \; & w^i(t) - D^i + \beta^i tb^i \hat{w}^{iB} \\ & + bm^{Rii} tb^{Ri} \hat{w}^{RiB} + m^{Rii} w^{Ri}(t) \\ & - m^{iRi} w^i(t) - md^{iRi} w^i(t) \end{aligned} \quad (48)$$

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

$$md^{iRi} = \frac{td^{Ri}}{1-td^{Ri}} \quad (49)$$

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 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. The sequence of operations involved can be followed on flow diagrams of the account building and of the projection procedures (Figure 6A and 6B).

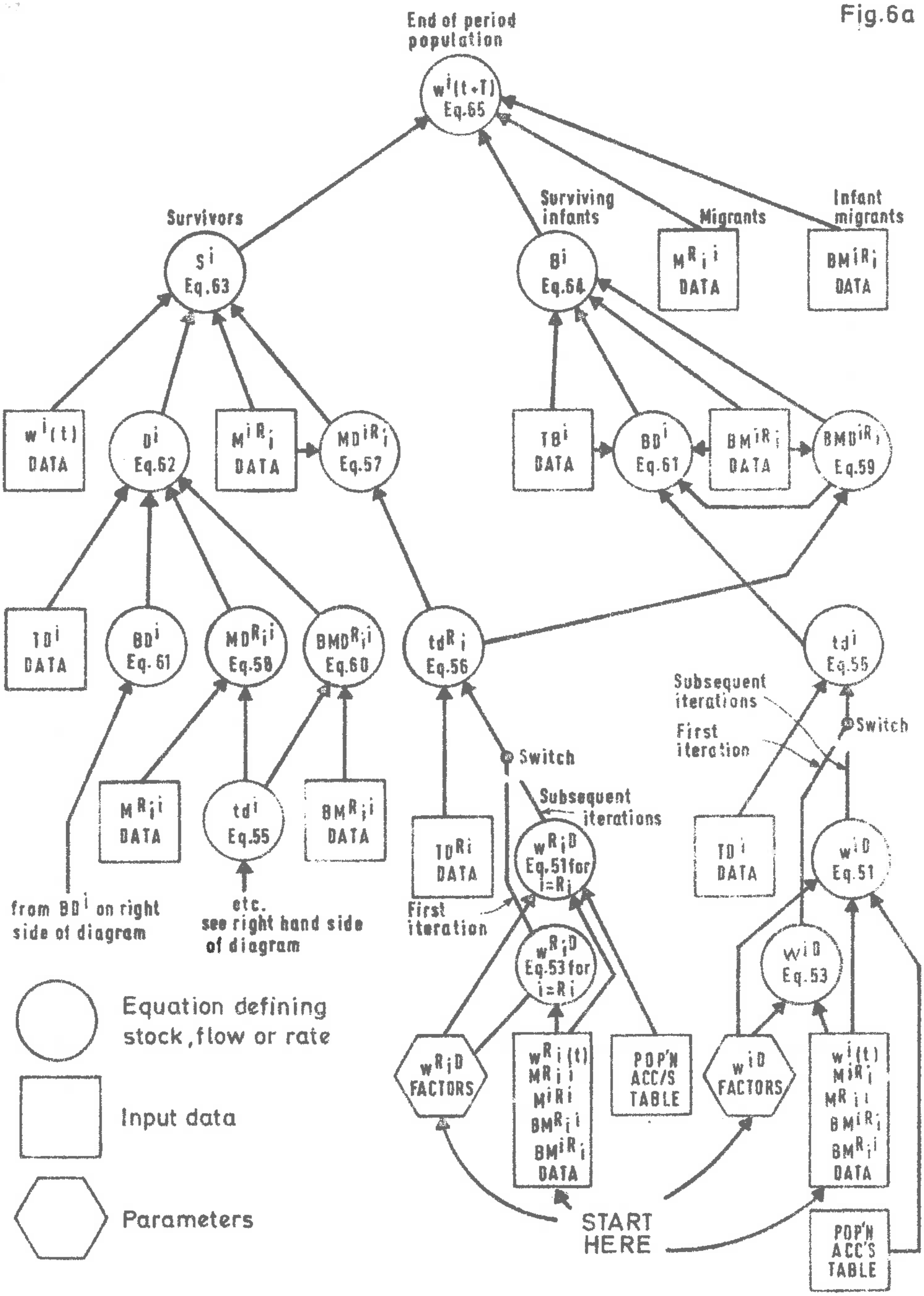
4.7.1 The accounts building procedures

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

(2) Calculate the at risk populations for births and for deaths:

$$\hat{w}^{iB} = w^i(t) - \tfrac{1}{2}D^i - \tfrac{1}{2}M^{iRi} - \tfrac{3}{4}MD^{iRi} + \tfrac{1}{2}M^{Rii} + \tfrac{1}{4}MD^{Rii} \quad (50)$$

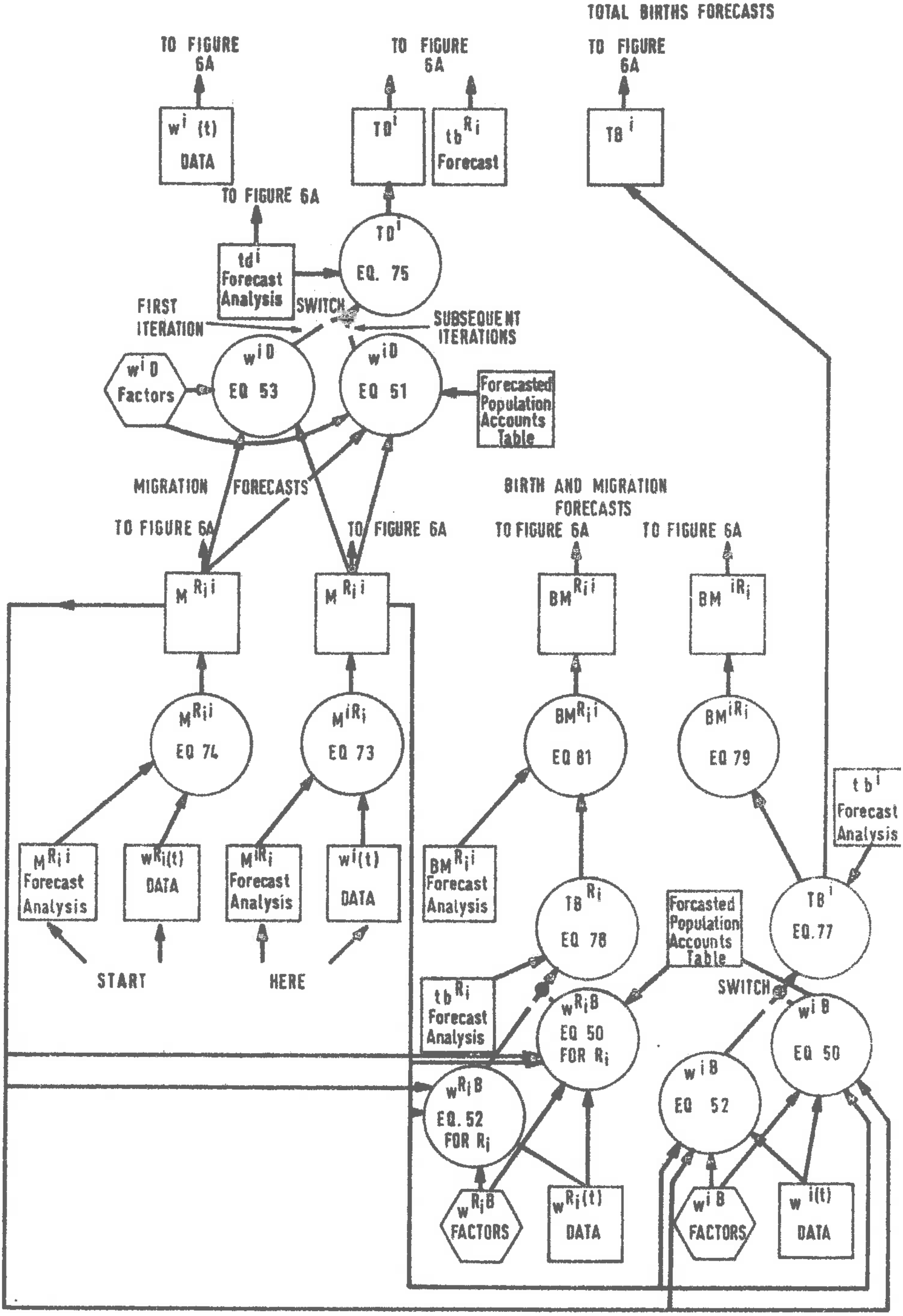
Fig.6a



The accounts building procedure

Fig. 6B

THE PROJECTION PROCEDURE



and

$$\hat{w}^{iD} = w^i(t)-\tfrac{1}{2}D^i-\tfrac{1}{2}M^{iR_i}-\tfrac{3}{4}MD^{iR_i}$$
$$+\tfrac{1}{2}M^{R_ii}+\tfrac{1}{4}MD^{R_ii}-\tfrac{1}{2}B^i-\tfrac{1}{4}BD^i$$
$$+\tfrac{1}{4}BM^{iR_i}+\tfrac{1}{8}BMD^{iR_i}+\tfrac{1}{4}BM^{R_ii}+\tfrac{1}{8}BMD^{iR_i} \qquad (51)$$

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

$$\hat{w}^{iB} = w^i(t)-\tfrac{1}{2}M^{iR_i}+\tfrac{1}{2}M^{R_ii} \qquad (52)$$

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

$$\text{and} \quad \hat{w}^{iD} = w^i(t)-\tfrac{1}{2}M^{iR_i}+\tfrac{1}{2}M^{R_ii}-\tfrac{1}{4}BM^{iR_i}+\tfrac{1}{4}BM^{R_ii} \qquad (53)$$

$$\text{or} \quad \hat{w}^{iD} = w^i(t) \qquad (54)$$

Equivalent equations are used to define $\hat{w}^{R_iB}$ and $\hat{w}^{R_iD}$. The population at risk of giving birth is not used directly in the accounting 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:

(3) Calculate td^i from

$$td^i = TD^i/\hat{w}^{iD} \qquad (55)$$

and td^{R_i} from

$$td^{R_i} = TD^{R_i}/\hat{w}^{R_iD} \qquad (56)$$

(4) Calculate MD^{iR_i}, MD^{R_ii}, BMD^{iR_i}, BMD^{R_ii}, and BD^i from the following equations

$$MD^{iR_i} = \frac{td^{R_i}\, M^{iR_i}}{1-td^{R_i}} \qquad (57)$$

$$MD^{R_i i} = \frac{td^i \; M^{R_i i}}{1-td^i} \tag{58}$$

$$BMD^{iR_i} = \frac{td^{R_i} \; BM^{iR_i}}{1-td^{R_i}} \tag{59}$$

$$BMD^{R_i i} = \frac{td^i \; BM^{R_i i}}{1-td^i} \tag{60}$$

$$BD^i = td^i(TB^i-BM^{iR_i}-BMD^{iR_i}) \tag{61}$$

with equivalent equations for the corresponding R_i terms.

(5) Once the smaller unknown flows have been calculated the larger unknown flows can be estimated:

$$D^i = TD^i-BD^i-MD^{R_i i}-BMD^{R_i i} \tag{62}$$

$$S^i = w^i(t)-D^i-M^{iR_i}-MD^{iR_i} \tag{63}$$

and

$$B^i = TB^i-BD^i-BM^{iR_i}-BMD^{iR_i} \tag{64}$$

The terms D^{R_i}, S^{R_i} and B^{R_i} can be calculated from the region R_i equivalents of (62), (63) and (64).

(6) End of period population may now be calculated using

$$w^i(t+T) = S^i+B^i+M^{R_i i}+BM^{R_i i} \tag{65}$$

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. These steps 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*

Migration rates are given by

$$m^{iR_i} = M^{iR_i}/w^i(t) \tag{66}$$

$$m^{R_ii} = M^{R_ii}/w^{R_i}(t) \tag{67}$$

Total death rates have been defined in equations (55) and (56).
Total Birth rates are obtained from

$$tb^i = TB^i/\hat{w}^{iB} \tag{68}$$

and

$$tb^{R_i} = TB^{R_i}/\hat{w}^{R_iB} \tag{69}$$

Birth and migration rates are defined as follows as

$$\beta m^{R_ii} = BM^{R_ii}/TB^{R_i} \tag{70}$$

and

$$\beta m^{iR_i} = BM^{iR_i}/TB^i \tag{71}$$

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

$$bm^{R_ii} = BM^{R_ii}/w^{R_i}(t) \tag{72}$$

and

$$bm^{iR_i} = BM^{iR_i}/w^i(t) \tag{73}$$

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

$$M^{iR_i} = m^{iR_i}w^i(t) \tag{74}$$

$$M^{R_ii} = m^{R_ii}w^{R_i}(t) \tag{75}$$

$$TD^i = td^i\hat{w}^{iD} \tag{76}$$

* Methods and models for doing this are discussed in later papers (Rees and Wilson 1972, and Wilson and Rees 1972B).

$$TD^{Ri} = td^{Ri}\hat{w}^{RiD} \qquad (77)$$

$$TB^{i} = tb^{i}\hat{w}^{iB} \qquad (78)$$

$$TB^{Ri} = tb^{Ri}\hat{w}^{RiB} \qquad (79)$$

$$\beta M^{iRi} = \beta m^{iRi}tb^{i}\hat{w}^{iB} \qquad (80)$$

or

$$BM^{iRi} = bm^{iRi}w^{i}(t) \qquad (81)$$

$$BM^{Rii} = \beta m^{Rii}tb^{Ri}\hat{w}^{RiB} \qquad (82)$$

or

$$BM^{Rii} = bm^{Rii}w^{Ri}(t) \qquad (83)$$

We can express the constituent terms of the $w^{i}(t+T)$ equation (equation (65)) in model form as follows:

$$S^{i} = w^{i}(t)-(td^{i}\hat{w}^{iD} -td^{i}(tb^{i}\hat{w}^{iB}-\beta m^{iRi}tb^{i}\hat{w}^{iB}-\frac{td^{Ri}}{1-td^{Ri}}\beta m^{iRi}tb^{i}\hat{w}^{iB}) -\frac{td^{i}}{1-td^{i}}m^{Rii}w^{Ri}(t) -\frac{td^{i}}{1-td^{i}}\beta m^{Rii}tb^{Ri}\hat{w}^{RiB}) -m^{iRi}w^{i}(t) -\frac{td^{Ri}}{1-td^{Ri}}m^{iRi}w^{i}(t) \qquad (84)$$

$$B^{i} = tb^{i}\hat{w}^{iB}-td^{i}(tb^{i}\hat{w}^{iB}-\beta m^{iRi}tb^{i}\hat{w}^{iB}-\frac{td^{Ri}}{1-td^{Ri}}\beta m^{iRi}tb^{i}\hat{w}^{iB}) -\beta m^{iRi}tb^{i}\hat{w}^{iB} -\frac{td^{Ri}}{1-td^{Ri}}\beta m^{iRi}tb^{i}\hat{w}^{iB} \qquad (85)$$

$$M^{R_i i} = m^{R_i i} w^{R_i}(t) \tag{86}$$

and

$$BM^{R_i i} = \beta m^{R_i i} tb^{R_i} \hat{w}^{R_i B} \tag{87}$$

Equations (84)-(87) can be combined to give a fully explicit model equation for $w^i(t+T)$. This equation differs from previous model equations (e.g. equation (40)) by including only rates calculated from known flows.

Equation (79) says that all that is nesessary 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}, 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.

(a) 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.

(b) Calculate the migration flow terms using equations (74) and (75). 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 (83) and (82). Otherwise the calculation of birth and migration terms is postponed to a later step.

(c) Calculate the at risk populations as in step (2) of the accounting building procedures setting the unestimated terms to zero. Initially, equation (52) will be used for population at risk of birth; at subsequent iterations equation (50) will be used. For population at risk of dying initially the following equation will be used

$$\hat{w}^{iD} = w^i(t) - \tfrac{1}{2}M^{iR_i} + \tfrac{1}{2}M^{R_i i} \tag{88}$$

and subsequently equation (51), unless the birth and migration terms have already been estimated. In that case, equation (53) will be used initially and (51) in subsequent iterations.

(d) Calculate TD^i and TD^{R_i} from equations (76) and (77), TB^i and TB^{R_i} from (78) and (79), and $BM^{R_i i}$ and BM^{iR_i} from (80) and (82) unless the birth and migration terms have already been calculated.

(e) 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.

(f) 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.

(g) The projected population can be calculated using equation (66) and the projected population accounts table is now complete for the first time.

(h) Cycle through steps (e)-(g) until convergence is achieved and a best estimate of the projected populations of region i and R_i are obtained.

5. A worked example

5.1 The account building calculations

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, U.K. for the census period 1961 to 1966, and how it can be used in projecting the population.

Figure 7 sets out the population accounts table in conceptual form and Figure 8 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 the Table. 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 in-migrate and survive there in 1966. These have to be estimated from data on the number of children in in-migrant families. All the migration figures have to be inflated by a correction factor to allow for the known under-enumeration of the Sample Census of 1966. Out-migration 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, 1972-B) 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 Figure 8 are the best estimates we are likely to get for these flows.

The problem is now 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 Figure 8, 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 out-migrants 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 ~~more~~ reflect 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 zone into a vastly larger one is likely to be more closely approximated

Figure 7 <u>The population accounts table for the West Riding of Yorkshire, 1961-1966 census period: conceptual terms</u>

Time → t+T / Region ↓ t		Yorkshire, West Riding Region i		Rest of England and Wales Region N_i		Rest of the World Region R		Totals Population Terms
		Alive in 1966	Died 1961-66	Alive in 1966	Died 1961-66	Alive in 1966	Died 1961-66	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 1961-66	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 1961-66	$BM^{N_i i}$	$BMD^{N_i i}$	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^{Ri}	MD^{Ri}	M^{RN_i}	ND^{RN_i}	S^R	D^R	$w^{R2}(t)$ $(w^{R1}(t))$
	Born 1961-66	BM^{Ri}	BMD^{Ri}	BM^{RN_i}	BMD^{RN_i}	B^R	BD^R	TB^R
Population Terms*								
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})		

*

<u>Notes</u>

* The column population totals refer to time t+T.

<u>Regions</u>

i Yorkshire, West Riding (1961 definition)

N_i Rest of England and Wales i.e. England and Wales minus Yorkshire, West Riding

R Rest of the World i.e. 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.)

Figure 8 The population accounts table for the West Riding of Yorkshire, 1961-1966 showing known flows

Time / Region	Region→ / Time (t, t+T)	Yorkshire, West Riding Region i		Rest of England and Wales Region Ni		Rest of the World Region R		Totals Population Terms
		Alive in 1966	Died 1961-66	Alive in 1966	Died 1961-66	Alive in 1966	Died 1961-66	Resident (Enumerated)
Yorkshire, West Riding Region i	Alive in 1961			168,207		55,328		3,650,586 (3,644,582)
	Born 1961-66			7,391		4,221		333,321
Rest of England and Wales Region N_i	Alive in 1961	137,183				831,390		42,453,962 (42,459,966)
	Born 1961-66	7,327				57,031		3,925,236
Rest of the World Region R	Alive in 1961	58,804		1,012,460		—	—	—
	Born 1961-66	2,509		44,358		—	—	—
Totals: Accounts total Inflated resident (Resident) (Enumerated)		3,741,153 (3,688,720) (3,686,500)	226,529	44,069,974 (43,452,330) (43,449,010)	2,537,801	—	—	—

by the smaller zone's rate than the larger, 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.

(1) death rates, first iteration

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:

$$td^k = TD^k/w^k(t) \tag{89}$$

for k = i,

$$\begin{aligned} td^i &= 226{,}529/3{,}650{,}586 = 0.062053 \\ &= 62.053/1000 \end{aligned} \tag{90}$$

and for $k = N_i$

$$\begin{aligned} td^{Ni} &= 2{,}537{,}801/42{,}453{,}962 = 0.059778 \\ &= 59.778/1000 \end{aligned} \tag{91}$$

(2) Migration and death factors, first iteration

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

$$\frac{td^k}{(1-td^k)} \tag{92}$$

for k = i

$$td^i/(1-td^i) \quad = \quad 0.066158 \quad = \quad 66.158/1000 \tag{93}$$

and

$$td^{Ni}(1-td^{Ni}) \quad = \quad 0.063578 \quad = \quad 63.578/1000 \tag{94}$$

These factors are then applied as follows:

$$MD^{iNi} \quad = \quad [td^{Ni}/(1-td^{Ni})]M^{iNi}$$

$$= \quad (0.63578)(168,207) \quad = \quad 10,694 \tag{95}$$

$$MD^{iR} \quad = \quad [td^i/(1-td^i)]M^{iR}$$

$$= \quad (0.066158)(55,328) \quad = \quad 3,660 \tag{96}$$

$$BMD^{iNi} \quad = \quad [td^{Ni}/(1-td^{Ni})]BM^{iNi}$$

$$= \quad (0.63578)(7,391) \quad = \quad 470 \tag{97}$$

and so on. Terms M^{iR}, BM^{iR}, M^{Nii}, BM^{Nii}, M^{Ri} and BM^{Ri} are multiplied by $|td^i/(1-td^i)|$ to yield MD^{iR}, BMD^{iR}, MD^{Nii}, BMD^{Nii}, MD^{Ri} and BMD^{Ri}. Terms M^{iNi}, BM^{iNi}, M^{NiR}, BM^{NiR}, M^{RNi}, and BM^{RNi} are multiplied by $|(td^{Ni}/(1-td^{Ni})|$ to yield the equivalent migration and death terms. Figure 9 shows the state of the accounts at the end of step 2. 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.

(3) <u>birth and death terms, first iteration</u>

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

$$BD^k \quad = \quad td^k(TB^k - \sum_{j\neq k} BM^{kj} - \sum_{j\neq k} BMD^{kj}) \tag{98}$$

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

Figure 9 The population accounts table for the West Riding of Yorkshire, 1961-66 at the end of step (2)

Time / Region	Region→ / Time; t+T / t	Yorkshire, West Riding Region i: Alive in 1966	Yorkshire, West Riding Region i: Died 1961-66	Rest of England and Wales Region N_i: Alive in 1966	Rest of England and Wales Region N_i: Died 1961-66	Rest of the World Region R: Alive in 1966	Rest of the World Region R: Died 1961-66	Totals Population terms Resident (Enumerated)
Yorkshire, West Riding Region i	Alive in 1961			168,207	10,694	55,328	3,660	3,650,586
	Born 1961-66			7,391	470	4,221	279	333,321
Rest of England and Wales Region N_i	Alive in 1961	137,183	9,076			831,390	52,858	42,453,962
	Born 1961-66	7,327	485			57,031	3,626	3,925,236
Rest of the World Region R	Alive in 1961	58,804	3,890	1,012,460	63,370	–	–	–
	Born 1961-66	2,509	166	44,358	2,820	–	–	–
Totals: Accounts total Inflated resident (Resident) (Enumerated)			226,529		2,537,801	–	–	–

For k = i equation (10) becomes

$$BD^{i} = td^{i}(TB^{i}-BM^{iN_i}-BMD^{iN_i}-BM^{iR}-BMD^{iR})$$

$$= (0.072053)(333{,}321-7{,}391-470-4{,}221-279)$$

$$= (0.062653)(320{,}960) = 19{,}916 \qquad (99)$$

For k = N_i equation (86) becomes

$$BD^{N_i} = TD^{N_i}(TB^{N_i}-BM^{N_i i}-BMD^{N_i i}-BM^{N_i R}-BMD^{N_i R})$$

$$= 0.059778(3{,}925{,}236-7{,}327-485-57{,}031-3{,}626)$$

$$= 0.059778(3{,}856{,}767)$$

$$= 230{,}547 \qquad (100)$$

(4) <u>deaths in situ terms, first iteration</u>

Having calculated the birth and death term, it is now possible to calculate the deaths in situ term. In general, this is given by the following relation

$$D^{k} = TD^{k}-BD^{k}-\sum_{j\neq k} MD^{kj}-\sum_{j\neq k} BMD^{kj} \qquad (101)$$

For k = i, equation (13) becomes

$$D^{i} = TD^{i}-BD^{i}-MD^{N_i i}-BMD^{N_i i}-MD^{Ri}-BMD^{Ri}$$

$$= 226{,}529-19{,}916-9{,}076-485-3{,}890-166$$

$$= 192{,}996 \qquad (102)$$

For the rest of England and Wales, region N_i, the deaths term is given by

$$D^{N_i} = TD^{N_i}-BD^{N_i}-MD^{iN_i}-BMD^{iN_i}-MD^{RN_i}-BMD^{RN_i}$$

$$= 2{,}537{,}801-230{,}547-10{,}694-570-64{,}370-2{,}820$$

$$= 2{,}228{,}900 \qquad (103)$$

(5) the survival terms, first iteration

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

$$S^k = w^k(t)-D^k-\sum_{j\neq k} M^{kj}-\sum_{j\neq k} MD^{kj} \tag{104}$$

For the West Riding, region i, this translates as

$$\begin{aligned} S^i &= w^i(t)-D^i-M^{iNi}-MD^{iNi}-M^{iR}-MD^{iR} \\ &= 3,650,586-192,996-168,207-10,794-55,328-3,660 \\ &= 3,219,701. \end{aligned} \tag{105}$$

For region N_i this becomes

$$\begin{aligned} S^{Ni} &= w^{Ni}(t)-D^{Ni}-M^{Nii}-MD^{Nii}-M^{NiR}-MD^{NiR} \\ &= 42,453,962-2,228,900-137,183-9,076-831,390-52,850 \\ &= 39,194,555 \end{aligned} \tag{106}$$

Surviving infants born in the period are calculated using

$$B^k = TB^k-BD^k-\sum_{j\neq k} BM^{kj}-\sum_{j\neq k} BMD^{kj} \tag{107}$$

which for West Yorkshire is

$$\begin{aligned} B^i &= TB^i-BD^i-BM^{iNi}-BMD^{iNi}-BM^{iR}-BMD^{iR} \\ &= 333,321-19,916-7,391-470-4,221-279 \\ &= 301,044 \end{aligned} \tag{108}$$

and for the Rest of England and Wales is

$$\begin{aligned} B^{Ni} &= TB^{Ni}-BD^{Ni}-BM^{Nii}-BMD^{Nii}-BM^{NiR}-BMD^{NiR} \\ &= 3,925,236-230,547-7,327-485-57,031-3,626 \\ &= 3,626,220 \end{aligned} \tag{109}$$

(6) calculation of the closing stock populations, first iteration

The final step in the sequence is the calculation of the new, end of period populations, using the general equation

$$w^k(t+T) \quad = \quad S^k + B^k + \sum_{j\neq k} M^{jk} + \sum_{j\neq k} BM^{jk} \tag{110}$$

In the West Riding case this is

$$w^i(t+T) \quad = \quad S^i + B^i + M^{Ni\,i} + BM^{Ni\,i} + M^{Ri} + BM^{Ri}$$

$$= \quad 3,219,701+301,044+137,183+7,327+58,804+2,509$$

$$= \quad 3,726,568 \tag{111}$$

and for the Rest of England and Wales we have the following version of (99)

$$w^{Ni}(t+T) \quad = \quad S^{Ni} + B^{Ni} + M^{iNi} + BM^{iNi} + M^{RNi} + BM^{RNi}$$

$$= \quad 39,194,555+2,626,220+168,207+7,391+1,002,460+44,358$$

$$= \quad 44,053,191 \tag{112}$$

Figure 10 shows the state of the accounts after the first six steps.

(7) calculation of populations at risk of dying, second iteration

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 w^{*iD} is calculated by adding up all the flow terms in the population accounting matrix weighted by the average proportion of the period that they are exposed in the region in question (sections 4.4-4.6). Table 3 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 lie 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.

Figure 10 <u>The population accounts table for the West Riding of Yorkshire, 1961-66 at the end of step (6)</u>

Region→ / Time ↓	Time t / t+T	Yorkshire, West Riding Region i: Alive in 1966	Yorkshire, West Riding Region i: Died 1961-66	Rest of England and Wales Region N_i: Alive in 1966	Rest of England and Wales Region N_i: Died 1961-66	Rest of the World Region R: Alive in 1966	Rest of the World Region R: Died 1961-66	Totals Population Terms Resident (Enumerated)
Yorkshire West Riding Region i	Alive in 1961	3,219,701	192,996	168,207	10,694	55,328	3,660	3,650,586
	Born 1961-66	301,044	19,916	7,391	470	4,221	279	333,321
Rest of England and Wales Region N_i	Alive in 1961	137,183	9,076	39,194,555	2,228,900	831,390	52,858	42,453,962
	Born 1961-66	7,327	485	3,626,220	230,547	57,031	3,626	3,925,236
Rest of the World Region R	Alive in 1961	58,804	3,890	1,012,460	63,370	—	—	—
	Born 1961-66	2,509	166	44,358	2,820	—	—	—
Totals: Accounts total		3,726,568		44,053,191				
Inflated resident (Resident) (Enumerated)			226,529		2,537,801	—	—	—

Table 3 Calculation of the populations at risk of dying, 1961-1966

Term (1)	Flow (2)	Fraction for region i (3)	Constituent of $\hat{w}^{iD}$ (4) = (2)(3)	Fraction for region Ni (5)	Constituent of $\hat{w}^{N_iD}$ (6) = (2)(5)
S^i	3,219,701	1	3,219,701	0	0
D^i	192,996	$\frac{1}{2}$	96,498	0	0
M^{iN_i}	168,207	$\frac{1}{2}$	84,103.5	$\frac{1}{2}$	84,103.5
MD^{iN_i}	10,694	$\frac{1}{4}$	2,673.5	$\frac{1}{4}$	2,673.5
M^{iR}	55,328	$\frac{1}{2}$	27,664	0	0
MD^{iR}	3,660	$\frac{1}{4}$	915	0	0
B^i	301,044	$\frac{1}{2}$	150,522	0	0
BD^i	19,916	$\frac{1}{4}$	4,979	0	
BM^{iN_i}	7,391	$\frac{1}{4}$	1,847.75	$\frac{1}{4}$	1,847.75
BMD^{iN_i}	470	$\frac{1}{8}$	58.75	$\frac{1}{8}$	58.75
BM^{iR}	4,221	$\frac{1}{4}$	1,055.25	0	0
BMD^{iR}	279	$\frac{1}{8}$	34.875	0	0
M^{N_ii}	137,183	$\frac{1}{2}$	68,591.5	$\frac{1}{2}$	68,591.5
MD^{N_ii}	9,076	$\frac{1}{4}$	2,269	$\frac{1}{4}$	2,269
S^{N_i}	39,194,555	0	0	1	39,194,555
D^{N_i}	2,228,900	0	0	$\frac{1}{2}$	1,114,450
M^{N_iR}	831,390	0	0	$\frac{1}{2}$	415,695
MD^{N_iR}	52,858	0	0	$\frac{1}{4}$	13,214.5
BM^{N_ii}	7,327	$\frac{1}{4}$	1,831.75	$\frac{1}{4}$	1,831.75
BMD^{N_ii}	485	$\frac{1}{8}$	60.625	$\frac{1}{8}$	60.625
B^{N_i}	3,626,220	0	0	$\frac{1}{2}$	1,813,110
BD^{N_i}	230,547	0	0	$\frac{1}{4}$	57,636.75
BM^{N_iR}	57,031	0	0	$\frac{1}{4}$	14,257.75
BMD^{N_iR}	3,626	0	0	$\frac{1}{8}$	453.25
M^{Ri}	58,804	$\frac{1}{2}$	29,402	0	0
MD^{Ri}	3,890	$\frac{1}{4}$	972.5	0	0
M^{RN_i}	1,012,460	0	0	$\frac{1}{2}$	506,230
MD^{RN_i}	63,370	0	0	$\frac{1}{4}$	16,092.5
S^R	-	0	0	0	0
D^R	-	0	0	0	0
BM^{Ri}	2,509	$\frac{1}{4}$	627.25	0	0
BMD^{Ri}	166	$\frac{1}{8}$	20.75	0	0
BM^{RN_i}	44,358	0	0	$\frac{1}{4}$	11,089.5
BMD^{RN_i}	2,820	0	0	$\frac{1}{8}$	352.5
B^R	-	0	0	0	0
BD^R	-	0	0	0	0
Totals			3,693,828		43,318,573.125

(8) death rates, second iteration

New death rates are calculated using $\hat{w}^{iD}$ and $\hat{w}^{NiD}$ as the base populations:

$$td^{i} = TD^{i}/\hat{w}^{iD}$$

$$= 226{,}529/3{,}693{,}828 = 0.061325$$

$$= 61.325/1000 \tag{113}$$

and

$$td^{Ni} = TD^{Ni}/\hat{w}^{NiD}$$

$$= 2{,}537{,}801/43{,}318{,}573.125 = 0.058585$$

$$= 58.585/1000 \tag{114}$$

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

(9) - (13), second iteration

Steps (9) to (13) repeat the calculations of steps (2) to (6) using the new death rates. The results of this second iteration are given in Figure 11 (the upper figures in the flow 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 Figure 11 is compared with the results of the first iteration (Figure 9) some small changes can be seen to have occurred. Each of terms that depend on the death rate directly is lower in Figure 11 (the MD, BMD and BD terms). The D^{i} and D^{Ni} terms are consequently larger and in situ survivors, S^{i} and S^{Ni}, 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 3,726,528 in the case of the West Riding and only 343 (decrease) out of 44,053,191 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.

5.2 Assessment of the performance of the account building method

Although the validity of the account building method described here depends on its own internal logic rather than empirical verification, it would be useful to compare our final population estimates with actual census counts of the population of the West Riding and the Rest of England and Wales.

Figure 11 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

Region (Time t)	Time t / t+T	Yorkshire, West Riding Region i: Alive in 1966	Yorkshire, West Riding Region i: Died 1961-66	Rest of England and Wales Region N_i: Alive in 1966	Rest of England and Wales Region N_i: Died 1961-66	Rest of the World Region R: Alive in 1966	Rest of the World Region R: Died 1961-66	Totals
Yorkshire, West Riding Region i	Alive in 1961	3,219,570 3,219,511	193,398 193,397	168,207	10,468 10,468	55,328	3,615 3,615	3,650,586
	Born 1961-66	301,289 301,289	19,684 19,684	7,391	460 460	4,221	276 276	333,321
Rest of England and Wales Region N_i	Alive in 1961	137,183	8,962 8,963	39,189,434 39,189,542	2,235,255 2,235,144	831,390	51,738 51,740	42,453,962
	Born 1961-66	7,327	479 479	3,630,998 3,630,890	225,852 225,960	57,031	3,549 3,549	3,925,236
Rest of the World Region R	Alive in 1961	58,804	3,842 3,842	1,012,460	63,006 63,008	-	-	-
	Born 1961-66	2,509	164 164	44,358	2,760 2,761	-	-	-
Totals: Accounts total		3,726,682 3,726,683		44,052,848 44,052,848		-		–
Inflated Resident		3,741,153				-		
(Resident)		3,668,720	226,529		2,537,801		-	
(Enumerated)		3,686,500						

Note

The upper figure in any cell represents the state of the PAT 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.

Figure 12 sets out the West Riding and Rest of England and Wales populations as estimated by the methods outlined in this paper and by a series of simpler, but conceptually invalid methods. The numbers under- or over-enumerated are given, and these are expressed as a percentage of the observed populations of the two regions.

The observed populations are themselves subject to error for two reasons. Firstly, the Sample Census of 1966 was discovered, as a result of post-enumeration checks, to have been an underenumeration of the population of England and Wales by some 670,000 (General Register Office, 1968, pp. xvi-vviii). Secondly, the population concept that should most probably be used in the accounting table is that of "estimated resident population (estimate definition)". This, however, is not available in 1966, and instead the "estimated resident population (census definition)" has to be used. Birth, death and migration figures all correspond most closely to usually resident population (estimate definition). The observed population for 1966 given in Figures 8, 11 and 12 are the usually resident populations (census definition) inflated by an England and Wales factor of 1.0142143 to allow for under-enumeration. This introduces a third source of error possible in the estimated observed populations. The inflation factor needed may vary from region to region. No information is given by the General Register Office that would make possible the use of a West Riding inflation factor. It may well be, therefore, that our accounting method gives a better estimate of the usually resident population of the West Riding than does the inflated census count.

The first two population accounting or projection methods are included in Figure 12 only to serve as a base from which comparisons can be drawn. The first method involves the hypotheses that

$$w^i(t+T) = w^i(t) \tag{115}$$

and

$$w^{Ni}(t+T) = w^{Ni}(t) \tag{116}$$

These substantially underestimate the populations of the West Riding and the Rest of England and Wales. The method involving simple extrapolation using an exponential growth rate uses the following equation

$$w^i(t+T) = w^i(t)e^{rn} \tag{117}$$

where r is the rate of growth and n the number of years over which it applies.

Figure 12 A comparison of estimated and observed populations for the West Riding of Yorkshire and the Rest of England and Wales at census date 1966.

Source of Population Estimate	Yorkshire, West Riding			Rest of England and Wales		
	$w^{i}(t+T)$	Difference	%	$w^{Ni}(t+T)$	Difference	%
Estimated observed population, t+T	3,741,153	-	-	44,069,974	-	-
Stable population hypothesis	3,650,586	-90,567	-2.42	42,453,962	-1,616,012	-3.67
Extrapolation of 1951-61 exponential growth	3,683,000	-58,153	-1.55	43,740,000	-729,974	-1.66
Simple component method ignoring migration	3,757,378	+16,225	+0.43	43,841,397	-228,577	-0.52
Simple component method recognising migration	3,729,830	-11,323	-0.30	44,053,495	-16,479	-0.04
This paper	3,726,683	-14,470	-0.39	44,052,848	-17,036	-0.04

The rate of growth was given by

$$r = \frac{1}{10} \log e \frac{w^i(t)}{w^i(t-2T)} \qquad (118)$$

and n = 5. The error is less but still substantial.

The simple component methods are the ones usually used in population estimation. The first is based on the equation

$$w^i(t+T) = w^i(t)-TD^i+TB^i \qquad (119)$$

which ignores migration and suffers therefrom in terms of size of error. The second method recognises migration explicitly

$$w^i(t+T) = w^i(t)-TD^i+TB^i-\sum_{j\neq i} M^{ij}+\sum_{j\neq i} M^{ji} \qquad (120)$$

and produces estimates which are slightly better than those produced by the proper accounting methods. This does not mean that this is necessarily the best method but it does serve to show that an explicit concern with migration is an essential prerequisite of demographic accounting and population projection.

5.3 Use of the accounts in population projection

Here we describe how the West Riding's population may be projected using two methods. The first is a simple method that employs the rates or outflow coefficients calculated from the population accounts table. The second method is that outlined in the summary of model building procedures (section 4.7).

(1) Population accounts table rates, method 1

The first step is to calculate the outflow coefficients or rates associated with the population accounting matrix. This is done by dividing each flow in Figure 11 (third iteration figures) by its raw total. For example,

$$s^i = = S^i/w^i(t) = 3,219,571/3,650,586 = 0.8819 \qquad (121)$$

$$d^i = D^i/w^i(t) = 193,397/3,650,586 = 0.00530 \qquad (122)$$

.

.

$$\dot{b}^i = B^i/TB^i = 301,289/333,321 = 0.9039 \qquad (123)$$

$$bd^{i} = BD^{i}/TB^{i} = 19,684/333,321 = 0.0591 \quad (124)$$

.

.

$$m^{Ni}i = M^{Nii}/w^{Ni}(t) = 137,183/42,453,962 = 0.0032 \quad (125)$$

.

.

.

The coefficients for the rest of the world region have, however, to be calculated differently since we do not know the region R population nor region R total births. The region R coefficients are defined as follows:

$$m^{Ri} = M^{Ri}/w^{i}(t) = 58,804/3,650,586 = 0.0161 \quad (126)$$

$$md^{Ri} = MD^{Ri}/w^{i}(t) = 3,842/3,650,586 = 0.0011 \quad (127)$$

$$m^{RNi} = M^{RNi}/w^{Ni}(t) = 1,012,460/42,453,962 = 0.0238 \quad (128)$$

$$md^{RNi} = MD^{RNi}/w^{Ni}(t) = 63,008/42,453,962 = 0.0015 \quad (129)$$

$$bm^{Ri} = BM^{Ri}/TB^{i} = 2,509/333,321 = 0,0075 \quad (130)$$

$$bmd^{Ri} = BMD^{Ri}/TB^{i} = 164/333,321 = 0.0005 \quad (131)$$

$$bm^{RNi} = BM^{RNi}/TB^{Ni} = 44,358/3,925,236 = 0.0113 \quad (132)$$

$$bmd^{RNi} = BMD^{RNi}/TB^{Ni} = 2,761/3,425,236 = 0.0007 \quad (133)$$

The resulting outflow coefficients or rates are shown in Figure 13.

(2) <u>rate projections, method 1</u>

The second step is to project these outflow coefficients for the future time periods for which it is desired to generate population accounting matrices and closing stock populations. This will be attempted in a future paper, and in this paper we make the crude assumption that the coefficients remain stable for next time period, 1966 to 1971.

(3) <u>flow calculations, method 1</u>

The third step involves multiplying the opening stock populations of the regions fully recognised in the spatial system under investigation by the appropriate outflow coefficients projected for the future period. In the West Riding and Rest of England and Wales case this involves multiplying

Figure 13 Outflow coefficients for the West Riding and the Rest of England and Wales, 1961-1966

Time Region \ Region Time	t \ t+T	Yorkshire, West Riding Region i		Rest of England and Wales Region N_i		Rest of the World Region R		Totals
		Alive in 1966	Died 1961-66	Alive in 1966	Died 1961-66	Alive in 1966	Died 1961-66	
Yorkshire, West Riding Region i	Alive in 1961	0.8819	0.0530	0.0461	0.0029	0.0152	0.0010	1.0000
	Born 1961-66	0.9039	0.0591	0.0222	0.0014	0.0127	0.0008	1.0000
Rest of England a and Wales Region N_i	Alive in 1961	0.0032	0.0002	0.9231	0.0526	0.0196	0.0012	1.0000
	Born 1961-66	0.0019	0.0001	0.9250	0.0576	0.0145	0.0009	1.0000
Rest of the World Region R	Alive in 1961	0.0161	0.0011	0.0238	0.0015	-	-	-
	Born 1961-66	0.0075	0.0005	0.0113	0.0007	-	-	-
Totals		-	-	-	-	-	-	-

Note

The outflow coefficients in the Region R are not true outflow coefficients but are formed by dividing the appropriate flow by the time t population or period (t,t+T) births total of the destination region.

$$\text{e.g.} \quad m^{Ri} = M^{Ri}/w^{i}(t)$$

$$= 58{,}804/3{,}650{,}586$$

$$= 0.0161$$

the Figure 13 outflow coefficients in the first, third and firth rows by the numbers in the first, third and fifth column totals cell of Figure 11. The resulting projected flows are given in Figure 14.

(4) total births, method 1

Before we can apply the projected outflow coefficients for the accounting matrix's rows involving birth, we have to project the total numbers likely to be born in the fully recognised regions in the projection period. Two alternatives are available. We can either multiply the opening stock population by an opening stock defined birth rate or we can multiply the population at risk of giving birth over the projection period by a population at risk defined birth rate. The first model, written formally is

$$TB^{i}(t+T,t+2T) \quad = \quad tb^{i}(1)w^{i}(t+T) \tag{134}$$

where under the assumption made in step two

$$tb^{i}(1) \quad = \quad TB^{i}(t,t+T)/w^{i}(t) \tag{135}$$

and

$TB^{i}(t+T,t+2T)$ = projected total births in region i in the projection period (t+T,t+2T), census period 1966-1971.

The second model is:

$$TB^{i}(t+T,t+2T) \quad = \quad tb^{i}(2)\hat{w}^{iB}(t+T,t+2T) \tag{136}$$

where

$$tb^{i}(2) \quad = \quad TB^{i}(t,t+T)/\hat{w}^{iB}(t,t+T) \tag{137}$$

Projected births using equation (134) are given for the West Riding by

$$TB^{i}(t+T,t+2T) \quad = \quad (333{,}321/3{,}650{,}586)3{,}726{,}683$$

$$= \quad (0.091306)3{,}726{,}683 \quad = \quad 340{,}269 \tag{138}$$

and for the Rest of England and Wales by

$$TB^{Ni}(t+T,t+2T) \quad = \quad (3{,}925{,}236/42{,}453{,}962)44{,}052{,}848$$

$$= \quad (0.092549)(44{,}052{,}848) \quad = \quad 4{,}073{,}067 \tag{139}$$

Figure 14 Projected population accounts table for the West Riding of Yorkshire and the Rest of England and Wales, 1966-71, after step three of the first projection method

Time / Region	Region / Time (t \ t+T)	Yorkshire, West Riding Region i: Alive in 1961	Yorkshire, West Riding Region i: Died 1966-71	Rest of England and Wales Region N_i: Alive in 1971	Rest of England and Wales Region N_i: Died 1966-71	Rest of the World Region R: Alive in 1971	Rest of the World Region R: Died 1966-71	Totals
Yorkshire, West Riding Region i	Alive in 1966	3,286,683	197,428	171,713	10,686	56,481	3,690	3,726,683
	Born 1966-71							
Rest of England and Wales Region N_i	Alive in 1966	142,350	9,300	40,665,484	2,319,323	862,701	53,689	44,052,848
	Born 1966-71							
Rest of the World Region R	Alive in 1966	59,929	3,922	1,050,591	65,381	-	-	-
	Born 1966-71							
Totals								

The second births model involves calculating the population at risk of giving birth. This is set out in Table 4 using the definitions of population at risk of giving birth set out in section 4.5. Predicted births in the West Riding in 1966-1971 are therefore

$$TB^{i}(t+T,t+2T) = (333{,}321/3{,}532{,}752.5)3{,}607{,}533$$

$$= (0.0943)3{,}607{,}533 = 340{,}269 \qquad (140)$$

Total births predicted in the Rest of England and Wales are

$$TN^{N_i}(t+T,t+2T) = (3{,}925{,}236/41{,}415{,}278.75)42{,}973{,}587$$

$$= (0.0948)42{,}973{,}587 = 4{,}073{,}067 \qquad (141)$$

The two models produce the same predicted births totals but only because of the assumption of constant coefficients. If some change had been assumed in the coefficients then the two predictions would have been different. The second model is more satisfying from a conceptual point of view; however, it is probably more difficult to project the population at risk birth rate than the opening stock population birth rate.

(5) births, method 1

The next step in the projection procedure is to multiply the predicted births by the appropriate coefficients in the second, fourth and sixth rows of Figure 13. This then fills in the population accounting matrix.

(6) closing stock populations, method 1

The final step, as with the historical accounts, is the calculation of the closing stock populations. The region i column terms and the region N_i column terms are totalled to yield predicted 1971 populations. The complete projected accounts table for 1966-1971 is shown in Figure 15 together with the population figures available from preliminary reports from the 1971 census. Assessment of the results of the projection is postponed until after the second method has been described.

We now proceed to method 2.

(1) analysis of historical rates, method 2

The second method uses the same body of information as the first but uses in a different way which may prove to be more convenient. The same

Table 4 Calculation of the populations at risk of giving birth, 1961-1966 and 1966-1971

Period (1)	Term (2)	Flow* (3)	Fraction for region i (4)	Constituent of $\hat{w}^{iB}$ (5) = (3)(4)	Fraction for region N_i (6)	Constituent of $\hat{w}^{N_iB}$ (7) = (3)(6)
(t,t+T)	S^i	3,219,571	1	3,219,571	0	0
or	D^i	193,397	$\frac{1}{2}$	96,698.50	0	0
1961-66	M^{iN_i}	168,207	$\frac{1}{2}$	84,103.50	$\frac{1}{2}$	84,103.50
	MD^{iN_i}	10,468	$\frac{1}{4}$	2,617	$\frac{1}{4}$	2,617
	M^{iR}	55,328	$\frac{1}{2}$	27,664	0	0
	MD^{iR}	3,615	$\frac{1}{4}$	903.75	0	0
	$M^{N_i i}$	137,183	$\frac{1}{2}$	68,591.50	$\frac{1}{2}$	68,591.50
	$MD^{N_i i}$	8,963	$\frac{1}{4}$	2,240.75	$\frac{1}{4}$	2,240.75
	S^{N_i}	39,189,542	0	0	1	39,189,542
	D^{N_i}	2,235,144	0	0	$\frac{1}{2}$	1,117,572
	$M^{N_i R}$	831,390	0	0	$\frac{1}{2}$	415,695
	$MD^{N_i R}$	51,740	0	0	$\frac{1}{4}$	12,935
	$M^{R i}$	58,804	$\frac{1}{2}$	29,402	0	0
	$MD^{R i}$	3,842	$\frac{1}{4}$	960.50	0	0
	$M^{R N_i}$	1,012,460	0	0	$\frac{1}{2}$	506,230
	$MD^{R N_i}$	63,008	0	0	$\frac{1}{4}$	15,752
	S^R	-	0	0	0	0
	D^R	-	0	0	0	0
	Totals			3,532,752.50		41,415,278.75
(t+T,t+2T)	S^i	3,286,683	1	3,286,683	0	0
or	D^i	197,428	$\frac{1}{2}$	98,714	0	0
1966-71	M^{iN_i}	171,713	$\frac{1}{2}$	85,856.5	$\frac{1}{2}$	85,856.50
	MD^{iN_i}	10,686	$\frac{1}{4}$	2,671.5	$\frac{1}{4}$	2,671.50
	M^{iR}	56,481	$\frac{1}{2}$	28,240.5	0	0
	MD^{iR}	3,690	$\frac{1}{4}$	922.5	0	0
	$M^{N_i i}$	142,350	$\frac{1}{2}$	71,175.	$\frac{1}{2}$	71,175
	$MD^{N_i i}$	9,300	$\frac{1}{4}$	2,325	$\frac{1}{4}$	2,325
	S^{N_i}	40,665,484	0	0	1	40,665,484
	D^{N_i}	2,319,323	0	0	$\frac{1}{2}$	1,159,661.50
	$M^{N_i R}$	862,701	0	0	$\frac{1}{2}$	431,350.50
	$MD^{N_i R}$	53,689	0	0	$\frac{1}{4}$	13,422.25
	$M^{R i}$	59,929	$\frac{1}{2}$	29,964.5	0	0
	$MD^{R i}$	3,922	$\frac{1}{4}$	980.5	0	0
	$M^{R N_i}$	1,050,591	0	0	$\frac{1}{2}$	525,295.50
	$MD^{R N_i}$	65,381	0	0	$\frac{1}{4}$	16,345.25
	S^R	-	0	0	0	0
	D^R	-	0	0	0	0
	Totals			3,607,533.0		42,973,587.00

* From Figure 16, third iteration

Figure 15 Projected population accounts table for the West Riding of Yorkshire and the Rest of England and Wales, 1966-1971 after step six of the first projection method

Time Region (t) / Region Time (t+T)		Yorkshire, West Riding Region i: Alive in 1971	Yorkshire, West Riding Region i: Died 1966-71	Rest of England and Wales Region N_i: Alive in 1971	Rest of England and Wales Region N_i: Died 1966-71	Rest of the World Region R: Alive in 1971	Rest of the World Region R: Died 1966-71	Totals
Yorkshire, West Riding Region i	Alive in 1966	3,286,683	197,428	171,713	10,686	56,481	3,690	3,726,683
	Born 1966-71	307,569	20,094	7,545	470	4,309	282	340,269
Rest of England and Wales Region N_i	Alive in 1966	142,350	9,300	40,665,484	2,319,323	862,701	53,689	44,052,848
	Born 1966-71	7,603	497	3,767,635	234,470	59,179	3,683	4,073,067
Rest of the World Region R	Alive in 1966	59,929	3,922	1,050,591	65,381	-	-	-
	Born 1966-71	2,561	167	46,029	2,865	-	-	-
Totals: Accounts total		3,806,695	231,408	45,708,997	2,633,195	-	-	-
Corrected accounts		3,844,198		45,671,496				
Resident		3,783,465		44,813,119				
Enumerated		3,779,035		44,813,119				

strategy is adopted in projection as was adopted in constructing the accounts in the first place: the known flows are first projected and then the full accounting table is derived via iteration.

The first step with the second method is the same as with the first: an analysis of the known flow rates and event totals in the recent past. In our particular case, we calculate the migration, total birth and total death rates for the West Riding and the Rest of England and Wales for the period 1961-1977. The migration rates are as given in the outflow coefficient table (Figure 13), and the relevant birth and death rates are collected together in Figure 16.

(2) projection of the rates, method 2

The asuumption as before is made, that is, that the rates will remain stable over the two time periods 1961-1966 and 1966-1971.

(3) calculation of migration and survival terms, method 3

Opening stock populations are multiplied by the migration outflow coefficients only.

(4) projection of total births and deaths, method 2

Total births are projected as in step four of the first projection method using the first births model (equation (134)). Total deaths are projected using the equivalent deaths model:

$$\hat{TD}^{k} = td^{R}(1)w^{k}(t+T) \tag{142}$$

where

$$td^{k}(1) = TD^{k}(t,t+T)/w^{k}(t) \tag{143}$$

For the West Riding total deaths are given by

$$TD^{i} = (0.062053)(3,726,683)$$

$$= 231,252 \tag{144}$$

Rest of England and Wales deaths are

$$TD^{Ni} = (0.059778)(44,052,848)$$

$$= 2,633,391 \tag{145}$$

The projected accounts after step four are given in Figure 17 with the birth, migration and survival terms added.

Figure 16 Total birth and total death rates used in the projections using the second method

Rate (1)	Events Total (2)	Opening stock population (3)	Opening stock rate, $td^k(1)$ (4) = (2)/(3)	Population at risk (5)	Population at risk, rate $td^k(2)$ (6)=(2)/(3)
tb^i	333,321	3,650,586	0.091306	3,532,725.5	0.094352
tb^{Ni}	3,925,236	42,453,962	0.092459	41,415,278.75	0.094777
td^i	226,529	3,650,586	0.062053	3,693,852.325	0.061326
td^{Ni}	2,537,801	42,453,962	0.059778	43,317,119.750	0.058587

Figure 17 Projected population accounts table for the West Riding of Yorkshire and the Rest of England and Wales, 1966-1971, after step four of the second projection method.

Region	Time	Yorkshire, West Riding Region i		Rest of England and Wales Region N_i		Rest of the World Region R		Totals
Time / Region	t+T / t	Alive in 1971	Died 1966-71	Alive in 1971	Died 1966-71	Alive in 1971	Died 1966-71	
Yorkshire, West Riding Region i	Alive in 1966			171,713		56,481		3,726,683
	Born 1966-71			7,545		4,309		340,269
Rest of England and Wales Region N_i	Alive in 1966	142,350				862,701		44,052,848
	Born 1966-71	7,603				59,179		4,073,067
Rest of the World Region R	Alive in 1966	59,929		1,050,591		-	-	-
	Born 1966-71	2,561		46,029		-	-	-
Totals			231,252		2,633,391	-	-	-

(5)-(11) flow calculations, method 2

The next six steps in projecting the accounts table parellel the basic six steps described earlier for the historical accounts (sections 6.1 to 6.6). The results are shown in Figure 18. They differ very little from those in Figure 15 in closing stock populations, though more persons migrate and then die than in Figure 15.

(12) iteration, method 2

The crude version of the population accounts table now available can be used to generate populations at risk of dying and at risk of giving birth for 1966-1971. These would be multiplied by the projected $tb^{i}(2)$, $tb^{Ni}(2)$, $td^{i}(2)$ and $td^{Ni}(2)$ or population at risk rates to yield new births and deaths totals. We could then iterate again through the table using the usual procedures. This is not done here as the assumption of constant birth and death rates means that the same totals would be projected as are given in Figure 18.

5.4 Assessment of the projections

The projections of Figure 18 are adjusted for boundary changes and are compared with the latest figures available on the 1971 populations of the West Riding and the Rest of England and Wales in Figure 19. The errors are substantial and much larger than those for the historical accounts, as one might expect. The populations have been substantially over-predicted in the case of both regions.

The overpredictions undoubtedly stem from an overprediction of the 1966-1971 birth rate, and an overprediction of the in-migration rate into both regions from the rest of the world. These overpredictions stem from the assumption of stable rates. The lesson to be learnt is that careful attention in any predictions of regional population must be devoted to the projection of those rates (outflow coefficients or vital rates).

Figure 18 Projected population accounts table for the West Riding of Yorkshire and the Rest of England and Wales, 1966-1971, after the first and second iterations of the second projection

Region / Time	Time / Region	Yorkshire, West Riding Region i		Rest of England and Wales Region N_i		Rest of the World Region R		Totals
Time / Region	t+T / t	Alive in 1971	Died 1966-71	Alive in 1971	Died 1966-71	Alive in 1971	Died 1966-71	
Yorkshire, West Riding Region i	Alive in 1966	3,286,948	196,864	171,713	10,917	56,481	3,760	3,726,683
	Born 1966-71	307,317	20,333	7,545	480	4,309	285	340,269
Rest of England and Wales Region N_i	Alive in 1966	142,350	9,418	40,670,489	2,313,041	862,701	54,849	44,052,848
	Born 1966-71	7,603	503	3,762,788	239,232	59,179	3,762	4,073,067
Rest of the World Region R	Alive in 1966	59,929	3,965	1,050,591	66,795	-	-	-
	Born 1966-71	2,561	169	46,029	2,926	-	-	-
Totals		3,806,708	231,252	45,709,155	2,633,391	-	-	-

Figure 19 A comparison of projected and observed populations for the West Riding of Yorkshire and the Rest of England and Wales at census date 1971

Source of population estimate	Yorkshire, West Riding			Rest of England and Wales		
	$w^{i}(t+2T)$	Difference	%	$w^{Ni}(t+T)$	Difference	%
Enumerated population, 1971 census[a]	3,779,035	-	-	44,813,119	-	-
Estimated resident population, 1971[b]	3,783,465	-	-	44,810,193	-	-
Stable population hypothesis[c]	3,764,183	-19,282	-0.51	44,015,345	-794,848	-1.77
This paper, method two population	3,844,211	+60,746	+1.61	45,671,652	+861,459	+1.92

Notes

[a] The England and Wales figure is taken from the Census 1971, England and Wales, Preliminary Report, H.M.S.O., 1972. The Yorkshire, West Riding figure is taken from Census, 1971, Advance Analysis, Yorkshire, West Riding, H.M.S.O., 1972.

[b] The figures given in the sources noted under (a) are for enumerated populations. In 1961 and 1966 resident population for West Riding was an average of 4,430 (approx.) more than the enumerated, and the 1971 is adjusted accordingly by adding on 4,430. The England and Wales resident population is assumed to be the same as the enumerated, and the estimated resident population for the Rest of England and Wales only reflects the West Riding adjustment.

[c] These all include allowances for the territorial expansion of the West Riding between 1966 and 1971. This was assumed to add 37,503 onto the West Riding population projected from a 1966 base, and to reduce the Rest of England and Wales population by an equivalent amount.

References

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

General Register Office (1968) Sample Census 1966. England and Wales. Migration Regional Report. Yorkshire and Humberside Region. London, H.M.S.O.

Peacock, K., Wilson, A.G. and Rees, P.H. (1972) Models of migration, Working Paper in preparation, Department of Geograpny, University of Leeds.

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 8, Department of Geography, University of Leeds.

Rees, P.H. (1972-A) Measuring and modelling the distribution of social groups within cities. Working Paper 15, Department of Geography, University of Leeds. To be published in Alan Wilson (ed.) Patterns and processes in urban and regional systems, London, Pion.

Rees, P.H. (1972-B) Data and related estimation problems for spatial demographic analysis in the U.K. Working Paper in preparation, Department of Geography, University of Leeds.

Rees, P.H. and Wilson, A.G. (1972) The time and space framework of population accounting and projection. Working Paper in preparation, Department of Geography, University of Leeds.

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

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

Stone, R. (1971) Demographic accounting and model building, O.E.C.D. Technical Reports in Education and Development, Paris.

Wilson, A.G. (1971) Multi-regional models of population structure and some implications for a dynamic-residential location model. Working Paper 6, Department of Geography, University of Leeds. Forthcoming in Alan Wilson (ed.) Patterns and processes in urban and regional systems, London, Pion.

Wilson, A.G. and Rees, P.H. (1972-A) Accounts and models for spatial demographic analysis II: age/sex disaggregated populations. Working Paper in preparation, Department of Geography, University of Leeds.

Wilson, A.G. and Rees, P.H. (1972-B) Models of regional birth and death rates. Working Paper in preparation, Department of Geography, University of Leeds.

Appendix 1 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. The equations may be generalized very easily for a multi-regional system. 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

$$\hat{w}^{iB} = w^i(t)-\tfrac{1}{2}D^i-\tfrac{1}{2}\sum_{j\neq i} M^{ij}-\tfrac{3}{4}\sum_{j\neq i} MD^{ij}+\tfrac{1}{2}\sum_{j\neq i} M^{ji}+\tfrac{1}{4}\sum_{j\neq i} MD^{ji} \tag{146}$$

$$\hat{w}^{iD} = w^i(t)-\tfrac{1}{2}D^i-\tfrac{1}{2}\sum_{j\neq i} M^{ij}-\tfrac{3}{4}\sum_{j\neq i} MD^{ij}+\tfrac{1}{2}\sum_{j\neq i} M^{ji}+\tfrac{1}{4}\sum_{j\neq i} MD^{ji}-\tfrac{1}{2}B^i-\tfrac{1}{4}BD^i$$

$$+\tfrac{1}{4}\sum_{j\neq i} BM^{ij}+\tfrac{1}{8}\sum_{j\neq i} BMD^{ij}+\tfrac{1}{4}\sum_{j\neq i} BM^{ji}+\tfrac{1}{8}\sum_{j\neq i} BMD^{ij} \tag{147}$$

Total death rate

$$td^i = TD^i/\hat{w}^{iD} \quad \text{as before} \tag{148}$$

Migrants who die

$$MD^{ij} = \frac{td^j}{1-td^i} M^{ij} \tag{149}$$

Migrating infants who die

$$BMD^{ij} = \frac{td^j}{1-td^j} BM^{ij} \tag{150}$$

Infants who die in situ

$$BD^i = td^i(TB^i-\sum_{j\neq i} BM^{ij}-\sum_{j\neq i} BMD^{ij}) \tag{151}$$

Persons who die

$$D^i = TD^i-BD^i-\sum_{j\neq i} MD^{ji}-\sum_{j\neq i} BMD^{ji} \tag{152}$$

Survivors

$$S^i = w^i(t) - D^i - \sum_{j\neq i} M^{ij} - \sum_{j\neq i} MD^{ij} \tag{153}$$

Surviving infants

$$B^i = TB^i - BD^i - \sum_{j\neq i} BM^{ij} - \sum_{j\neq i} BMD^{ij} \tag{154}$$

End of period population

$$w^i(t+T) = S^i + B^i + \sum_{j\neq i} M^{ji} + \sum_{j\neq i} BM^{ji} \tag{155}$$

Migration rates

$$m^{ij} = M^{ij}/w^i(t) \tag{156}$$

Total birth rates

$$tb^i = TB^i/\hat{w}^{iB}(t) \quad \text{as before} \tag{157}$$

Infant migration rates

$$\beta m^{ij} = BM^{ij}/TB^i \tag{158}$$

Model equation for end of period population

$$\begin{aligned}
w^i(t+T) = {} & w^i(t) - \Big(td^i\hat{w}^{iD} - td^i\Big(tb^i\hat{w}^{iB} - \sum_{j\neq i} \beta m^{ij} tb^i\hat{w}^{iB} - \sum_{j\neq i} \frac{td^j}{1-td^j}\beta m^{ij} tb^i\hat{w}^{iB}\Big) \\
& - \sum_{j\neq i} \frac{td^i}{1-td^i} m^{ji} w^i(t) - \sum_{j\neq i} \frac{td^i}{1-td^i} \beta m^{ji} tb^j\hat{w}^{jB}\Big) - \sum_{j\neq i} m^{ij} w^i(t) \\
& - \sum_{j\neq i} \frac{td^j}{1-td^j} m^{ij} w^i(t) + tb^i\hat{w}^{iB} - td^i\Big(tb^i\hat{w}^{iB} - \sum_{j\neq i} \beta m^{ij} tb^i\hat{w}^{iB} \\
& - \sum_{j\neq i} \frac{td^j}{1-td^j} \beta m^{ij} tb^i\hat{w}^{iB}\Big) - \sum_{j\neq i} \beta m^{ij} tb^i\hat{w}^{iB} - \sum_{j\neq i} \frac{td^j}{1-td^j} \beta m^{ij} tb^i\hat{w}^{iB} \\
& + \sum_{j\neq i} m^{ji} w^j(t) + \sum_{j\neq i} \beta m^{ji} tb^j\hat{w}^{jB}
\end{aligned} \tag{159}$$

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 (144) to (159) then have to be modified. The modifications are left to the reader faced with a problem in population accounting and modelling.