Environment and Planning A, 1974, volume 6, pages 101-116

Accounts and models for spatial demographic analysis 2: age-sex disaggregated populations

A G Wilson, P H Rees
Department of Geography, University of Leeds, Leeds, England
Received 20 November 1973

Abstract. In an earlier paper, methods were presented for the construction of accounts and models for spatial demographic analysis. In this paper we show how these methods can be extended to deal with age-sex disaggregated populations. At the outset a new notation is introduced which facilitates disaggregation. The accounts are then defined and the basic accounting equations constructed. The appropriate rates and 'at-risk' populations are introduced as a preliminary to developing the equations of the account-based model. The model is also presented in transition-rate form, which facilitates comparison with alternative models.

1 Introduction: principles of disaggregation

In an earlier paper (Rees and Wilson, 1973a) we described how to build an account-based multiregional demographic model for populations in which age and sex were not distinguished. This allowed us to introduce a number of new concepts in a reasonably simple situation, but clearly a version of the model is needed which is applicable to age-sex disaggregated populations. In the first instance we concentrate in this paper on age disaggregation and assume that we are dealing with single-sex populations, say of females. At the end of the paper we indicate how sex disaggregation can be accomplished in a straightforward manner. In another paper (Rees, 1973) it was shown how a more convenient notation could be developed. In this paper we use this revised notation. We begin in this section by showing how it can be extended to deal with ageing, and the rest of the paper has a similar structure to our first paper. In section 2 we identify the basic accounting relationships, and in section 3 we develop the concept of 'at-risk populations'. This allows us to build the account-based model in section 4, and we investigate the transition-rate form of this model in section 5. There is a note on sex disaggregation in section 6 and some concluding comments are presented in section 7.

The treatment of ageing depends very much on the assumptions which are made about age groups and the length of the discrete time period, the so-called projection period, which is the basis of the analysis. There is one particularly simple case which arises when all the age group intervals are of equal length, each equal to the projection period. In this case any survivor 'ages' into the next highest age group during the projection period. In general the age group intervals can bear any relationship to each other and to the projection period, and no such simple ageing structure holds. We shall refer to the two cases as the 'simple' case and the 'general' case respectively. The relationship between them has been extensively explored in an earlier paper (Wilson, 1972). The simple case assumption has been much used by Rogers (1966, 1968).

The types of demographic flow introduced in our earlier paper (Rees and Wilson, 1973a) in the revised notation are represented as follows:

K^{ii}	nonmigrating survivors	$K^{\beta(i)i}$	nonmigrating surviving births
$K^{ij}, \quad i \neq j$	migrating survivors	$K^{\beta(i)j}, \quad i \neq j$	migrating surviving births
$K^{i\delta(i)}$	nonmigrating nonsurvivors	$K^{\beta(i)\delta(i)}$	nonmigrating nonsurviving births
$K^{i\delta(j)}, \quad i \neq j$	migrating nonsurvivors	$K^{\beta(i)\delta(j)}, \quad i \neq j$	migrating nonsurviving births.

The reader can easily make the transition to the old notation, if this is required, and record these flows in an accounting framework. We introduce age into each of these flows as follows.

K^{ii} can be subdivided into a set of flows K^{ii}_{rs}, where r is the age group at the initial time t, and s the age group at $t+T$. In the simple case only, such terms are nonzero if $s=r+1$, but in general there are more nonzero terms. Thus the K^{ii} element becomes a matrix $\mathbf{K}^{ii}$ whose (r,s)th element is K^{ii}_{rs}. If we assume that age groups are mutually exclusive and are numbered in order of increasing age, K^{ii}_{rs} can only be nonzero if $s \geqslant r$, and so the form of the matrix is

$$\mathbf{K}^{ii} = \begin{bmatrix} K^{ii}_{11} & K^{ii}_{12} & \ldots & K^{ii}_{1R-1} & K^{ii}_{1R} \\ 0 & K^{ii}_{22} & \ldots & K^{ii}_{2R-1} & K^{ii}_{2R} \\ 0 & 0 & \ldots & K^{ii}_{3R-1} & K^{ii}_{3R} \\ \cdot & \cdot & \cdots & & \cdot \\ \cdot & \cdot & \cdots & & \cdot \\ \cdot & \cdot & \cdots & & \cdot \\ 0 & 0 & \ldots & 0 & K^{ii}_{RR} \end{bmatrix}, \tag{1}$$

if there is a total of R age groups. Some of the elements which have been shown as nonzero may in fact also turn out to be zero—this depends on the age-group definitions in relation to the projection period.

The migration flows, K^{ij}, $i \neq j$, similarly become a matrix $\mathbf{K}^{ij}$, or $\{K^{ij}_{rs}\}$, whose (r,s)th element is K^{ij}_{rs}. In the case of deaths we also proceed similarly, but now take s as the age group at time of death and not as the age group at the end of the period. $K^{i\delta(i)}$ and $K^{i\delta(j)}$, $i \neq j$, become the matrices $\mathbf{K}^{i\delta(i)}$, $\mathbf{K}^{i\delta(j)}$, $i \neq j$, whose (r,s)th elements are $K^{i\delta(i)}_{rs}$ and $K^{i\delta(j)}_{rs}$ respectively.

The remaining four elements of the aggregated case, listed above, all involve births. In this case the first age group is unnecessary since it represents pre-existence. However, we need to record births by age of mother and it turns out to be sensible and convenient to use the first index to record this. Thus $K^{\beta(i)i}$ becomes the matrix $\mathbf{K}^{\beta(i)i}$ whose (r,s)th element is $K^{\beta(i)i}_{rs}$, and this is the number of nonmigrating surviving births in region i, to mothers in age group r at time t, recorded in age group s (for the child) at time $t+T$. The form of this matrix is

$$\mathbf{K}^{\beta(i)i} = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ \cdot & \cdot & \cdots & & \cdot \\ K^{\beta(i)i}_{\lambda 1} & K^{\beta(i)i}_{\lambda 2} & \ldots & K^{\beta(i)i}_{\lambda R-1} & K^{\beta(i)i}_{\lambda R} \\ \cdot & \cdot & \cdots & & \cdot \\ K^{\beta(i)i}_{\mu 1} & K^{\beta(i)i}_{\mu 2} & \ldots & K^{\beta(i)i}_{\mu R-1} & K^{\beta(i)i}_{\mu R} \\ \cdot & \cdot & \cdots & & \cdot \\ 0 & 0 & \ldots & 0 & 0 \end{bmatrix}, \tag{2}$$

where λ and μ are the lower and upper limits of the potential child-bearing age groups(1). The other three birth terms disaggregate similarly, though for $\mathbf{K}^{\beta(i)\delta(i)}$ and $\mathbf{K}^{\beta(i)\delta(j)}$, $i \neq j$, the second subscript, s, is age at time of death.

We have chosen to regard (r,s) as defining elements of a matrix which is itself labelled (i,i), (i,j) or whatever. We could also have regarded (i,r) as the definition of the initial state, or $[\beta(i),r]$, and (j,s), or $[\delta(j),s]$, as the definition of the final state. A variable such as K^{ij}_{rs} then represents the number of transitions between times t and $t+T$ from (i,r) to (j,s). From this viewpoint it will be seen that disaggregated accounts have exactly the same *structure* as aggregated ones.

(1) The lower and upper limits λ and μ refer to the age of mother at the beginning of the period, not the age of mother at time of birth. We have also assumed that $\lambda \geqslant 1$, that is, that female children born within the period do not themselves become mothers in the period.

2 Basic accounting relationships

We begin by making a distinction between different types of age group. Age groups 1, 2, ..., R are exhibited in relation to the projection period T for three different cases in figure 1. In any of the cases, births during t to $t+T$ could be noted as surviving into any of the age groups 1, 2, ..., m at time $t+T$. In case (a) there is no survival *into* age groups $s \leqslant m$; in cases (b) and (c) survival into age group m is possible, but not into $s < m$. Cases (b) and (c) are distinguished because survival within m is possible in case (c), but not in case (b). This subdivision of age groups, $s < m$, $s = m$, $s > m$ turns out to be significant for the accounting equations, as we shall see shortly.

These results can be summarised as follows. For case (a), $K_{rs}^{\beta(i)i}$ is nonzero for $\lambda \leqslant r \leqslant \mu$ and $s \leqslant m$, and survival terms such as K_{rs}^{ii} or K_{rs}^{ij}, $i \neq j$, are nonzero only for $s > m$. For cases (b) and (c), K_{rs}^{ii} may be nonzero for $s = m$ also, and for case (c), K_{mm}^{ii} may also be nonzero.

The accounting matrix, $\mathbf{K}$, for the age-disaggregated case may be constructed from the submatrices defined earlier as follows:

$$\mathbf{K} = \left[\begin{array}{cccc|cccc}
\mathbf{K}^{11} & \mathbf{K}^{12} & \ldots & \mathbf{K}^{1N} & \mathbf{K}^{1\delta(1)} & \mathbf{K}^{1\delta(2)} & \ldots & \mathbf{K}^{1\delta(N)} \\
\mathbf{K}^{21} & \mathbf{K}^{22} & \ldots & \mathbf{K}^{2N} & \mathbf{K}^{2\delta(1)} & \mathbf{K}^{2\delta(2)} & \ldots & \mathbf{K}^{2\delta(N)} \\
\cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\
\mathbf{K}^{N1} & \mathbf{K}^{N2} & \ldots & \mathbf{K}^{NN} & \mathbf{K}^{N\delta(1)} & \mathbf{K}^{N\delta(2)} & \ldots & \mathbf{K}^{N\delta(N)} \\
\hline
\mathbf{K}^{\beta(1)1} & \mathbf{K}^{\beta(1)2} & \ldots & \mathbf{K}^{\beta(1)N} & \mathbf{K}^{\beta(1)\delta(1)} & \mathbf{K}^{\beta(1)\delta(2)} & \ldots & \mathbf{K}^{\beta(1)\delta(N)} \\
\mathbf{K}^{\beta(2)1} & \mathbf{K}^{\beta(2)2} & \ldots & \mathbf{K}^{\beta(2)N} & \mathbf{K}^{\beta(2)\delta(1)} & \mathbf{K}^{\beta(2)\delta(2)} & \ldots & \mathbf{K}^{\beta(2)\delta(N)} \\
\cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\
\mathbf{K}^{\beta(N)1} & \mathbf{K}^{\beta(N)2} & \ldots & \mathbf{K}^{\beta(N)N} & \mathbf{K}^{\beta(N)\delta(1)} & \mathbf{K}^{\beta(N)\delta(2)} & \ldots & \mathbf{K}^{\beta(N)\delta(N)}
\end{array}\right] \tag{3}$$

for an N region system. The matrix could be written out in full if required by using the definitions of the submatrices in section 1, and the reader may like to do this for himself on a large sheet of paper!

There are two kinds of row sum:

$$K_{r*}^{i*} = \sum_j \sum_s K_{rs}^{ij} + \sum_j \sum_s K_{rs}^{i\delta(j)} \tag{4}$$

is the number of people alive in the (i, r) state at time t, and

$$K_{r*}^{\beta(i)*} = \sum_j \sum_s K_{rs}^{\beta(i)j} + \sum_j \sum_s K_{rs}^{\beta(i)\delta(j)} \tag{5}$$

is the number of births to mothers aged r at time t during the period t to $t+T$.

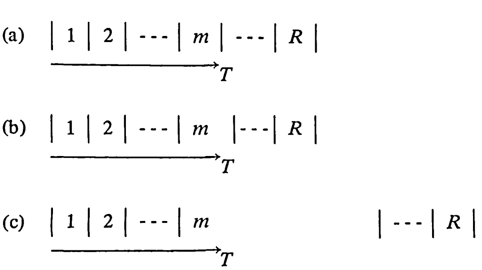

Figure 1. Possible age-group structures in relation to projection periods.

Similarly there are two kinds of column sum:

$$K_{*s}^{*j} = \sum_i \sum_r K_{rs}^{ij} + \sum_i \sum_r K_{rs}^{\beta(i)j} \tag{6}$$

is the number of people alive in the (j, s) state at time $t+T$, and

$$K_{*s}^{*\delta(j)} = \sum_i \sum_r K_{rs}^{i\delta(j)} + \sum_i \sum_r K_{rs}^{\beta(i)\delta(j)} \tag{7}$$

is the number of recorded deaths in region j during period t to $t+T$ of persons aged s at time of death.

In a formal sense, then, equations (4)–(7) represent the accounting equations for the age-disaggregated case for a single-sex population. Care has to be exercised in practice, at least to achieve computational economy, in carrying out the summations as we have not explicitly identified the nonzero elements. We can in fact write out the accounting equations in a more useful form by using the results obtained at the beginning of the section in which we differentiated the first m age groups from the rest. We also distinguish, for later convenience, '(i, i) terms' from '(i, j), $i \neq j$ terms'.

Equation (4) can now be written

$$\sum_{s \geqslant \max(r,m)} K_{rs}^{ii} + \sum_{j \neq i} \sum_{s \geqslant \max(r,m)} K_{rs}^{ij} + \sum_{s \geqslant r} K_{rs}^{i\delta(i)} + \sum_{j \neq i} \sum_{s \geqslant r} K_{rs}^{i\delta(j)} = K_{r*}^{i*} \, . \tag{8}$$

The notation '$\max(r, m)$' means 'the value of r or m, whichever is the greater'. Equation (5) can be written as

$$\sum_{s \leqslant m} K_{rs}^{\beta(i)i} + \sum_{j \neq i} \sum_{s \leqslant m} K_{rs}^{\beta(i)j} + \sum_{s \leqslant m} K_{rs}^{\beta(i)\delta(i)} + \sum_{j \neq i} \sum_{s \leqslant m} K_{rs}^{\beta(i)\delta(j)} = K_{r*}^{\beta(i)*} \, , \qquad \lambda \leqslant r \leqslant \mu \, . \tag{9}$$

We now write the 'total deaths' equations, (7), separately for $s \leqslant m$ and for $s > m$:

$$\sum_{r \leqslant s} K_{rs}^{i\delta(i)} + \sum_{j \neq i} \sum_{r \leqslant s} K_{rs}^{j\delta(i)} + \sum_{r=\lambda}^{\mu} K_{rs}^{\beta(i)\delta(i)} + \sum_{j \neq i} \sum_{r=\lambda}^{\mu} K_{rs}^{\beta(j)\delta(i)} = K_{*s}^{*\delta(i)} \, , \qquad s \leqslant m \, ; \tag{10}$$

and for $s > m$, there are no birth terms, and so

$$\sum_{r \leqslant s} K_{rs}^{i\delta(i)} + \sum_{j \neq i} \sum_{r \leqslant s} K_{rs}^{j\delta(i)} = K_{*s}^{*\delta(i)} \, , \qquad s > m \, . \tag{11}$$

In the new population equations, (6), we now distinguish three cases, $s < m$, $s = m$, and $s > m$:

$$\sum_{r=\lambda}^{\mu} K_{rs}^{\beta(i)i} + \sum_{j \neq i} \sum_{r=\lambda}^{\mu} K_{rs}^{\beta(j)i} = K_{*s}^{*i} \, , \qquad s < m \, , \tag{12}$$

$$\sum_{r \leqslant s} K_{rs}^{ii} + \sum_{j \neq i} \sum_{r \leqslant s} K_{rs}^{ji} + \sum_{r=\lambda}^{\mu} K_{rs}^{\beta(i)i} + \sum_{j \neq i} \sum_{r=\lambda}^{\mu} K_{rs}^{\beta(j)i} = K_{*s}^{*i} \, , \qquad s = m \, , \tag{13}$$

and

$$\sum_{r \leqslant s} K_{rs}^{ii} + \sum_{j \neq i} \sum_{r \leqslant s} K_{rs}^{ij} = K_{*s}^{*i} \, , \qquad s > m \, . \tag{14}$$

Note that if case (a) of figure 1 holds, then the equation containing births and those containing survivors are completely separated, since the survival flows in equation (13) are zero, and so equation (12) holds for $s \leqslant m$, though there is no corresponding simplification for deaths in equations (10) and (11). It would be a great convenience, therefore, to ensure that this does hold in age-group definitions.

In the simple case these accounting equations take a particularly convenient form. We would have $m = 1$, and survival terms would be K_{rs}^{ii} or K_{rs}^{ij} for $s = r+1$ (except for $r = R$, $s = R$), and the only nonzero death terms would require $s = r$ or $r+1$.

3 Rates and at-risk populations

In the development of an account-based age-disaggregated model we shall need to employ a number of birth and death rates. As with the aggregated case in our earlier paper, we shall define these in such a way as to utilise available birth and death information in the numerator, and this means that we must construct corresponding 'at-risk populations' as denominators. An 'at-risk population', therefore, is that set of people who are 'at-risk' during the period for the events referred to in the numerator.

In the case of births, we must construct an appropriate notation with some care. Births involve the characteristics of the mothers as well as the children. Let $K_{ru}^{i\mathrm{m}(j)}$ be the number of mothers who are in region i, age group r, at time t, and who give birth (that is, experience maternity) in region j and are in age group u at the time of the birth. We can use this quantity to *label* available birth data, which then appears as the (rather cumbersome) notation $K_{**}^{\beta(i)*}(K_{*u}^{*\mathrm{m}(i)})$. Let $\hat{K}_{*u}^{\mathrm{B}*i}$ be the at-risk population associated with this number of births, and then we can define a corresponding rate

$$b_{*u}^{*i} = \frac{K_{**}^{\beta(i)*}(K_{*u}^{*\mathrm{m}(i)})}{\hat{K}_{*u}^{\mathrm{B}*i}} . \tag{15}$$

What we should *like to have*, in that it connects more directly with the accounts, is a rate

$$b_{r*}^{i*} = \frac{K_{r*}^{\beta(i)*}(K_{r*}^{*\mathrm{m}(i)})}{\hat{K}_{r*}^{\mathrm{B}*i}} , \tag{16}$$

using a now obvious notation on the right hand side. We shall show later how to connect these two sets of rates.

In the case of deaths, the available information is $K_{*s}^{*\delta(i)}$, and this does match the accounts. The corresponding at-risk population is $\hat{K}_{*s}^{\mathrm{D}*i}$, and the rate is

$$d_{*s}^{*i} = \frac{K_{*s}^{*\delta(i)}}{\hat{K}_{*s}^{\mathrm{D}*i}} . \tag{17}$$

It is also convenient to assume we are given deaths by age at time t, $K_{rs}^{*\delta(i)}$; and if $\hat{K}_{rs}^{\mathrm{D}*i}$ is the corresponding at-risk population, we can also define

$$d_{rs}^{*i} = \frac{K_{rs}^{*\delta(i)}}{\hat{K}_{rs}^{\mathrm{D}*i}} . \tag{18}$$

We will use $r = 0$ to designate deaths of people born during the period, and shall show later how $K_{rs}^{*\delta(i)}$ can be obtained from $K_{*s}^{*\delta(i)}$.

We now indicate how at-risk populations are calculated, considering $\hat{K}_{r*}^{\mathrm{B}*i}$, $\hat{K}_{rs}^{\mathrm{D}*i}$, $\hat{K}_{*s}^{\mathrm{D}*i}$, and $\hat{K}_{*u}^{\mathrm{B}*i}$ in turn, this order being chosen for later convenience.

Our first task then is to disaggregate by age equation (34) from our earlier paper (Rees and Wilson, 1973a, p.73). There we defined some θ-coefficients which represented the proportions of time in which a given population flow was at-risk for the event. Formally, we can write

$$\hat{K}_{r*}^{\mathrm{B}*i} = \sum_{jk} {}^{i}\theta^{\mathrm{B}jk} K_{r*}^{jk} , \tag{19}$$

where ${}^{i}\theta^{\mathrm{B}jk}$ is a number, between 0 and 1, which is the proportion of the period that the K_{r*}^{jk} flow is at-risk for giving birth in i. The (j, k) summation is over all states. If we assume the θ's (as implied by our notation) to be age independent, then approximate values can be obtained from table 1 of our earlier paper (Rees and Wilson, 1973a, p.73) if the reader makes appropriate changes in notation. The fact that a formula such as that in equation (19) exists and is appropriate will suffice for the time being.

We can define another set of θ's, where we now distinguish age groups s, r, u corresponding to those shown on the flow variables, to obtain

$$\hat{K}_{rs}^{\mathrm{D}*i} = \delta_{r0} \sum_{j \in \mathrm{B}, k, u} {}_{s}^{i}\theta_{0u}^{\mathrm{D}jk} K_{*u}^{jk} + \sum_{j \in \mathrm{NB}, k, u} {}_{s}^{i}\theta_{ru}^{\mathrm{D}jk} K_{ru}^{jk} , \tag{20}$$

where the k summation is over all states, and $j \in \mathrm{B}$ denotes birth states, $j \in \mathrm{NB}$ nonbirth states. δ_{r0} is a Kronecker delta, which shows that the first set of terms contributes to $\hat{K}_{rs}^{\mathrm{D}*i}$ when $r = 0$. The θ's in this case are age dependent, and the procedure for calculating them is a fairly complicated one. Thus we restrict the presentation here to equation (20) and note that the full details are available elsewhere (Rees and Wilson, 1974). $\hat{K}_{*s}^{\mathrm{D}*i}$ can then be obtained as

$$\hat{K}_{*s}^{\mathrm{D}*i} = \sum_{r} \hat{K}_{rs}^{\mathrm{D}*i} \tag{21}$$

$$= \sum_{j \in \mathrm{B}, k, u} {}_{s}^{i}\theta_{0u}^{\mathrm{D}jk} K_{*u}^{jk} + \sum_{j \in \mathrm{NB}, k, u, r \neq 0} {}_{s}^{i}\theta_{ru}^{\mathrm{D}jk} K_{ru}^{jk} . \tag{22}$$

Finally we note that a similar technique can be used for $\hat{K}_{*u}^{\mathrm{B}*i}$: the same θ-weights can be used as in equation (22), except the first group of terms are omitted since we assume that T is sufficiently short that no-one can give birth who were themselves born in the period. Thus

$$\hat{K}_{*u}^{\mathrm{B}*i} = \sum_{j \in \mathrm{NB}, k, s, r \neq 0} {}_{u}^{i}\theta_{rs}^{\mathrm{D}jk} K_{rs}^{jk} . \tag{23}$$

4 The account-based model

We are now in a position to begin to assemble the account-based model for the age-disaggregated case. The basis of the model, of course, is the set of accounting equations (8)–(14). As in the aggregated case we have to give a procedure for estimating unknown minor flows, but now there is an additional difficulty: we still do not have enough equations to estimate all the major flows. We should expect to use equation (8) to give K_{rs}^{ii}, but there are R^2 variables and only R equations for each region. There are similar difficulties with $K_{rs}^{\beta(i)i}$ from equation (9) and $K_{rs}^{i\delta(i)}$ in equations (10) and (11). First, then, we show how to reduce the number of 'major flow' variables in the equations by defining various sets of coefficients, and then how to estimate the minor flows.

Define p_{rs} to be the proportion of survivors who were in age group r at time t, and who are in age group s at time $t+T$. We assume this to be true for any 'survivor' flow, and, in practice, our model calculation procedure builds this assumption into the model predictions. We would then have

$$K_{rs}^{ii} = p_{rs} K_{r*}^{ii} . \tag{24}$$

We can now substitute for K_{rs}^{ii} in equation (8) by using equation (24) and take K_{r*}^{ii} as the major flow variable. We then have exactly as many equations as unknowns, we can solve for K_{r*}^{ii} at the appropriate time, and then find K_{rs}^{ii} from equation (24).

For a historical analysis, we would assume migration flows of the form K_{rs}^{ij}, $i \neq j$ to be given, but actually we would have K_{*s}^{ij}, since this is the form in which migration data is available. This problem can be resolved by defining q_{rs} to be the proportion of people in age group s at time $t+T$ who were initially in age group r. Then, using these coefficients,

$$K_{rs}^{ij} = q_{rs} K_{*s}^{ij} . \tag{25}$$

In the births case we are given $K_{**}^{\beta(i)*}(K_{*u}^{*\mathrm{m}(i)})$. Let π_s be the proportion of these births which are recorded as being into age group s. Then the relationship

$$K_{*s}^{\beta(i)*}(K_{*u}^{*\mathrm{m}(i)}) = \pi_s K_{**}^{\beta(i)*}(K_{*u}^{*\mathrm{m}(i)}) \tag{26}$$

provides the appropriate disaggregation. The problem in equation (9) can be resolved by taking

$$K_{rs}^{\beta(i)i} = \pi_s K_{r*}^{\beta(i)i} , \tag{27}$$

by using the accounting equation to give $K_{r*}^{\beta(i)i}$, and then equation (27) to give $K_{rs}^{\beta(i)i}$.

In relation to deaths, we can define c_{rs} as the proportion of deaths in age groups of people who were initially in age group r. Then we can take

$$K_{rs}^{*\delta(i)} = c_{rs} K_{*s}^{*\delta(i)} . \tag{28}$$

This relationship can be substituted in equations (10) and (11), which are then used to obtain $K_{*s}^{*\delta(i)}$, and $K_{rs}^{*\delta(i)}$ is obtained from equation (28). A little thought shows that the same relationship can be used to obtain $K_{r*}^{\beta(i)*}$ from $K_{**}^{\beta(i)*}(K_{*u}^{*\mathrm{m}(i)})$,

$$K_{r*}^{\beta(i)*} = \sum_u c_{ru} K_{**}^{\beta(i)*}(K_{*u}^{*\mathrm{m}(i)}) . \tag{29}$$

One further coefficient is needed: e_{rs}, defined as the proportion of infant deaths associated with mothers aged r at time t. This would give, for example,

$$K_{rs}^{\beta(i)\delta(i)} = e_{rs} K_{*s}^{\beta(i)\delta(i)} . \tag{30}$$

We have thus defined five sets of coefficients: p_{rs}, q_{rs}, π_s, c_{rs}, and e_{rs}. Estimates of some of these can be made directly from the 'geometry' of the age group definitions in relation to the projection period, T. Others rely on the availability of data, say at the national scale. (In some cases, if the appropriate regional data exists, it may be possible to estimate coefficients on a regional basis and to add corresponding superscripts.) Results for p_{rs} and q_{rs} coefficients are given in an earlier paper (Wilson, 1972), where they arose in another context. A detailed account of methods for obtaining the full set of all coefficients, or of alternative, more detailed, procedures which fulfil the same role, is given in the book cited earlier (Rees and Wilson, 1974, especially chapters 9 and 17). In this paper we continue to concentrate on a presentation of the main concepts of the model, and henceforth we shall assume that values of the coefficients can be obtained for any particular case.

The next step in the argument concerns the problem of estimating the minor flows. In the aggregated case there were the $K^{i\delta(j)}$, $i \neq j$, $K^{\beta(i)\delta(j)}$, $i \neq j$, and $K^{\beta(i)\delta(i)}$ flows. In this case, we only need to consider the first two of these three types and we will give another procedure for the third.

We rely on the assumption used in our earlier paper that a rate is applicable to any subset of the corresponding set of events. Thus, to estimate $K_{rs}^{i\delta(j)}$, $i \neq j$, we assume that the age-specific death rate, d_{rs}^{*j}, is applicable to these deaths, and that

$$\frac{K_{rs}^{i\delta(j)}}{\hat{K}_{rs}^{\mathrm{D}i\delta(j)}} = d_{rs}^{*j} , \tag{31}$$

where $\hat{K}_{rs}^{\mathrm{D}i\delta(j)}$ is the at-risk population for this subset and is also a subset of the at-risk population $\hat{K}_{r*s}^{\mathrm{D}*j}$—defined, for region i, in equation (20). Picking out the relevant terms, we have

$$\hat{K}_{rs}^{\mathrm{D}i\delta(j)} = \sum_{u \geqslant s} {}_s^j\theta_{ru}^{\mathrm{D}ij} K_{ru}^{ij} + \sum_{u \geqslant s} {}_s^j\theta_{ru}^{\mathrm{D}i\delta(j)} K_{ru}^{i\delta(j)} . \tag{32}$$

We can substitute from equation (32) into equation (31) and solve for $K_{rs}^{i\delta(j)}$ to give

$$K_{rs}^{i\delta(j)} = \frac{d_{rs}^{*j}\left(\sum_{u \geqslant s} {}_s^j\theta_{ru}^{\mathrm{D}ij} K_{ru}^{ij} + \sum_{u > s} {}_s^j\theta_{ru}^{\mathrm{D}i\delta(j)} K_{ru}^{i\delta(j)}\right)}{1 - d_{rs}^{*j}\,{}_s^j\theta_{rs}^{\mathrm{D}i\delta(j)}} . \tag{33}$$

This equation is more difficult to handle than the equivalent one in the aggregate case because the right hand side still contains terms in $K_{ru}^{i\delta(j)}$, $u > s$. In other words, they are simultaneous equations in $K_{rs}^{i\delta(j)}$. They are *linear* simultaneous equations, and could be solved as such. However, since the $K_{ru}^{i\delta(j)}$ terms on the right hand side will be small relative to the K_{ru}^{ii} terms, it may be simpler to obtain an iterative solution: set the right hand side $K_{ru}^{i\delta(j)}$'s to zero, solve for $K_{rs}^{i\delta(j)}$, resubstitute and solve again, and so on.

We can use a similar method to obtain $K_{rs}^{\beta(i)\delta(j)}$, $i \neq j$. We should recall here that r is age of the mother at time t. The subset of the at-risk population of equation (20) which is relevant here may be designated $\hat{K}_{0s}^{\mathrm{D}\beta(i)\delta(j)}$, where the 0 subscript indicates a birth, and the corresponding death rate is d_{0s}^{*j}. We now further subdivide $\hat{K}_{0s}^{\mathrm{D}*j}$ by age of mother, r, and write it as ${}_r\hat{K}_{0s}^{\mathrm{D}\beta(i)\delta(j)}$. $\hat{K}_{0s}^{\mathrm{D}\beta(i)\delta(j)}$ may be obtained by picking out the relevant terms of equation (20) as

$$\hat{K}_{0s}^{\mathrm{D}\beta(i)\delta(j)} = \sum_{u \geqslant s} {}_s^j\theta_{0u}^{\mathrm{D}\beta(i)j} K_{*u}^{\beta(i)j} + \sum_{u \geqslant s} {}_s^j\theta_{0u}^{\mathrm{D}\beta(i)\delta(j)} K_{*u}^{\beta(i)\delta(j)} , \tag{34}$$

and ${}_r\hat{K}_{0s}^{\mathrm{D}\beta(i)\delta(j)}$ is obtained from this expression by replacing the $K_{*u}^{\beta(i)j}$ and $K_{*u}^{\beta(i)\delta(j)}$ terms by $K_{ru}^{\beta(i)j}$ and $K_{ru}^{\beta(i)\delta(j)}$ respectively, to give

$${}_r\hat{K}_{0s}^{\mathrm{D}\beta(i)\delta(j)} = \sum_{u \geqslant s} {}_s^j\theta_{0u}^{\mathrm{D}\beta(i)j} K_{ru}^{\beta(i)j} + \sum_{u \geqslant s} {}_s^j\theta_{0u}^{\mathrm{D}\beta(i)\delta(j)} K_{ru}^{\beta(i)\delta(j)} . \tag{35}$$

Then $K_{rs}^{\beta(i)\delta(j)}$ can be obtained from

$$\frac{K_{rs}^{\beta(i)\delta(j)}}{{}_r\hat{K}_{0s}^{\mathrm{D}\beta(i)\delta(j)}} = d_{0s}^{*j} . \tag{36}$$

We substitute from equation (35) into (36) and solve for $K_{rs}^{\beta(i)\delta(j)}$ to give

$$K_{rs}^{\beta(i)\delta(j)} = \frac{d_{0s}^{*j}\left(\sum\limits_{u \geqslant s} {}_s^j\theta_{0u}^{\mathrm{D}\beta(i)j} K_{ru}^{\beta(i)j} + \sum\limits_{u > s} {}_s^j\theta_{0u}^{\mathrm{D}\beta(i)\delta(j)} K_{ru}^{\beta(i)\delta(j)}\right)}{1 - d_{0s}^{*j}\, {}_s^j\theta_{rs}^{\mathrm{D}\beta(i)\delta(j)}} . \tag{37}$$

Once again these are linear simultaneous equations and we will assume, as before, that they are most conveniently solved iteratively. This completes our discussion on the estimation of minor flows.

We can now proceed to specify the account-based model for the age-disaggregated case. There are seven steps:

Step 1: assemble known data.
Step 2: manipulate this where necessary.
Step 3: obtain initial values of unknown major flows from the accounting equations, by setting unknown minor flows to zero.
Step 4: calculate at-risk populations.
Step 5: calculate birth and death rates.
Step 6: calculate unknown minor flows.
Step 7: solve the accounting equations for the unknown major flows.

Steps 4–7 must be repeated and the iteration continued until convergence is achieved. We shall now specify each step in detail using the arguments presented earlier in the paper. In each case, equations are presented in the form most convenient for computational purposes.

Step 1. The known data are assumed to consist of K_{r*}^{i*}; $K_{**}^{\beta(i)*}(K_{*u}^{*\mathrm{m}(i)})$; $K_{*s}^{*\delta(i)}$; K_{*s}^{ij}, $i \neq j$; $K_{rs}^{\beta(i)j}$, $i \neq j$.

Step 2. We can manipulate the birth data into a more convenient form using c-coefficients to give

$$K_{r*}^{\beta(i)*} = \sum_{u=\lambda}^{\mu} c_{ru} K_{**}^{\beta(i)*}(K_{*u}^{*\mathrm{m}(i)}) , \tag{38}$$

and we can disaggregate the migration data using q-coefficients as

$$K_{rs}^{ii} = q_{rs}K_{*s}^{ij} , \qquad i \neq j . \tag{39}$$

These quantities can now be treated as known 'data'.

Step 3. We next obtain initial values of the unknown major flows from the accounting equations, setting unknown minor flows to zero. We write the equations in their order of solution. We also assume that, where appropriate, terms such as $\sum\limits_{r \leqslant s} K_{rs}^{i\delta(i)}$ have been replaced by $K_{*s}^{i\delta(i)}$.

From equation (10) we obtain

$$K_{*s}^{i\delta(i)} + K_{*s}^{\beta(i)\delta(i)} = K_{*s}^{*\delta(i)} , \qquad s \leqslant m . \tag{40}$$

We can subdivide the terms on the left hand side by using the appropriate c-coefficients. Thus,

$$K_{*s}^{\beta(i)\delta(i)} = c_{0s}(K_{*s}^{i\delta(i)} + K_{*s}^{\beta(i)\delta(i)}) , \qquad s \leqslant m \tag{41}$$

and

$$K_{*s}^{i\delta(i)} = (1 - c_{0s})(K_{*s}^{i\delta(i)} + K_{*s}^{\beta(i)\delta(i)}) , \qquad s \leqslant m . \tag{42}$$

We can then use e-coefficients to disaggregate $K_{*s}^{\beta(i)\delta(i)}$ by age of mother:

$$K_{rs}^{\beta(i)\delta(i)} = e_{rs}K_{*s}^{\beta(i)\delta(i)} , \qquad s \leqslant m , \tag{43}$$

and use c-coefficients to obtain

$$K_{rs}^{i\delta(i)} = c_{rs}(K_{*s}^{i\delta(i)} + K_{*s}^{\beta(i)\delta(i)}) , \qquad s \leqslant m . \tag{44}$$

Equation (11) gives

$$K_{*s}^{i\delta(i)} = K_{*s}^{*\delta(i)} , \qquad s > m , \tag{45}$$

and use of c-coefficients gives

$$K_{rs}^{i\delta(i)} = c_{rs}K_{*s}^{i\delta(i)} , \qquad s > m . \tag{46}$$

Equation (9) can be solved for $K_{r*}^{\beta(i)i}$ as

$$K_{r*}^{\beta(i)i} = K_{r*}^{\beta(i)*} - \sum_{j \neq i}\sum_{s \leqslant m} K_{rs}^{\beta(i)j} - \sum_{s \leqslant m} K_{rs}^{\beta(i)\delta(i)} , \tag{47}$$

and then use of π-coefficients gives

$$K_{rs}^{\beta(i)i} = \pi_s K_{r*}^{\beta(i)i} . \tag{48}$$

Equation (8) gives

$$K_{r*}^{ii} = K_{r*}^{i*} - \sum_{j \neq i} K_{rs}^{ij} - \sum_{s \geqslant r} K_{rs}^{i\delta(i)} , \tag{49}$$

and p-coefficients give

$$K_{rs}^{ii} = p_{rs}K_{r*}^{ii} . \tag{50}$$

It is not necessary at this stage to calculate the new populations using equations (12)–(14).

Step 4. The at-risk populations needed are $\hat{K}_{r*}^{\mathrm{B}*i}$, $\hat{K}_{rs}^{\mathrm{D}*i}$, $\hat{K}_{*s}^{\mathrm{D}*i}$, and $\hat{K}_{*u}^{\mathrm{B}*i}$. These are obtained from equations (19), (20), (21), and (23), and are repeated here for convenience:

$$\hat{K}_{r*}^{\mathrm{B}*i} = \sum_{jk} {}^{i}\theta^{\mathrm{B}jk} K_{r*}^{jk} , \tag{51}$$

Calculation of at-risk populations (continued)

$$\hat{K}_{rs}^{\mathrm{D}*i} = \delta_{r0} \sum_{j\in \mathrm{B}, k, u} {}_{s}^{i}\theta_{0u}^{\mathrm{D}jk} K_{*u}^{jk} + \sum_{j\in \mathrm{NB}, k, u} {}_{s}^{i}\theta_{ru}^{\mathrm{D}jk} K_{ru}^{jk} , \tag{52}$$

$$\hat{K}_{*s}^{\mathrm{D}*i} = \sum_{r} \hat{K}_{rs}^{\mathrm{D}*i} \tag{53}$$

$$\hat{K}_{*u}^{\mathrm{B}*i} = \sum_{j\in \mathrm{NB}, k, s, r \neq 0} {}_{u}^{i}\theta_{rs}^{\mathrm{D}jk} K_{rs}^{jk} . \tag{54}$$

Step 5. The birth and death rates needed are b_{r*}^{*i}, b_{*s}^{*i}, d_{rs}^{*i}, and d_{*s}^{*i}. They are obtained from

$$b_{r*}^{*i} = \frac{K_{r*}^{\beta(i)*}}{\hat{K}_{r*}^{\mathrm{B}*i}} , \tag{55}$$

$$b_{*s}^{*i} = \frac{K_{**}^{\beta(i)*}(K_{*s}^{*\mathrm{m}(i)})}{\hat{K}_{*s}^{\mathrm{B}*i}} , \tag{56}$$

$$d_{rs}^{*i} = \frac{K_{rs}^{*\delta(i)}}{\hat{K}_{rs}^{\mathrm{D}*i}} , \tag{57}$$

$$d_{*s}^{*i} = \frac{K_{*s}^{*\delta(i)}}{\hat{K}_{*s}^{\mathrm{D}*i}} . \tag{58}$$

Step 6. The unknown minor flows are $K_{rs}^{i\delta(j)}$, $i \neq j$, and $K_{rs}^{\beta(i)\delta(j)}$, $i \neq j$. They are obtained from equations (33) and (37) which are repeated here for convenience:

$$K_{rs}^{i\delta(j)} = \frac{d_{rs}^{*j}\left(\sum\limits_{u \geqslant s} {}_{s}^{j}\theta_{ru}^{\mathrm{D}ij} K_{ru}^{ij} + \sum\limits_{u > s} {}_{s}^{j}\theta_{ru}^{\mathrm{D}i\delta(j)} K_{ru}^{i\delta(j)}\right)}{1 - d_{rs}^{*j}\,{}_{s}^{j}\theta_{rs}^{\mathrm{D}i\delta(j)}} . \tag{59}$$

$$K_{rs}^{\beta(i)\delta(j)} = \frac{d_{0s}^{*j}\left(\sum\limits_{u \geqslant s} {}_{s}^{j}\theta_{0u}^{\mathrm{D}\beta(i)j} K_{ru}^{\beta(i)j} + \sum\limits_{u > s} {}_{s}^{j}\theta_{0u}^{\mathrm{D}\beta(i)\delta(j)} K_{ru}^{\beta(i)\delta(j)}\right)}{1 - d_{0s}^{*j}\,{}_{s}^{j}\theta_{rs}^{\mathrm{D}\beta(i)\delta(j)}} . \tag{60}$$

It should be recalled that each of these equations must be solved iteratively.

Step 7. We can now solve the accounting, and related, equations as in *step 3*, but now using all terms with the unknown flows as calculated in *step 6*. We proceed, following the pattern of *step 3*, as follows. Equation (9) gives

$$K_{*s}^{i\delta(i)} + K_{*s}^{\beta(i)\delta(i)} = K_{*s}^{*\delta(i)} - \sum_{j \neq i} \sum_{r \leqslant s} K_{rs}^{j\delta(i)} - \sum_{j \neq i} \sum_{r} K_{rs}^{\beta(j)\delta(i)} , \qquad s \leqslant m \tag{61}$$

and then c- and e-coefficients can be used to give

$$K_{*s}^{\beta(i)\delta(i)} = c_{0s}(K_{*s}^{i\delta(i)} + K_{*s}^{\beta(i)\delta(i)}) , \tag{62}$$

$$K_{*s}^{i\delta(i)} = (1 - c_{0s})(K_{*s}^{i\delta(i)} + K_{*s}^{\beta(i)\delta(i)}) , \tag{63}$$

$$K_{rs}^{\beta(i)\delta(i)} = e_{rs} K_{*s}^{\beta(i)\delta(i)} , \tag{64}$$

$$K_{rs}^{i\delta(i)} = c_{rs} K_{*s}^{i\delta(i)} . \tag{65}$$

Equation (11) gives

$$K_{*s}^{i\delta(i)} = K_{*s}^{*\delta(i)} - \sum_{j \neq i} \sum_{r \leqslant s} K_{rs}^{j\delta(i)} , \qquad s > m , \tag{66}$$

and c-coefficients give

$$K_{rs}^{i\delta(i)} = c_{rs} K_{*s}^{i\delta(i)} . \tag{67}$$

Equation (9) gives

$$K_{r*}^{\beta(i)i} = K_{r*}^{\beta(i)*} - \sum_{j \neq i} \sum_{s \leqslant m} K_{rs}^{\beta(i)j} - \sum_{s \leqslant m} K_{rs}^{\beta(i)\delta(i)} - \sum_{j \neq i} \sum_{s \leqslant m} K_{rs}^{\beta(i)\delta(j)} , \tag{68}$$

and π-coefficients give

$$K_{rs}^{\beta(i)i} = \pi_s K_{r*}^{\beta(i)i} . \tag{69}$$

Equation (8) gives

$$K_{r*}^{ii} = K_{r*}^{i*} - \sum_{j \neq i} \sum_{s \geqslant r} K_{rs}^{ij} - \sum_{s \geqslant r} K_{rs}^{i\delta(i)} - \sum_{j \neq i} \sum_{s \geqslant r} K_{rs}^{i\delta(j)} , \tag{70}$$

and p-coefficients give

$$K_{rs}^{ii} = p_{rs} K_{r*}^{ii} . \tag{71}$$

Finally we can solve equations (12)-(14) for the new population totals, K_{*s}^{*i}:

$$K_{*s}^{*i} = \sum_{r} K_{rs}^{\beta(i)i} + \sum_{j \neq i} \sum_{r} K_{rs}^{\beta(j)i} , \qquad s < m , \tag{72}$$

$$K_{*s}^{*i} = \sum_{r \leqslant s} K_{rs}^{ii} + \sum_{j \neq i} \sum_{r \leqslant s} K_{rs}^{ji} + \sum_{r} K_{rs}^{\beta(i)i} + \sum_{j \neq i} \sum_{r} K_{rs}^{\beta(j)i} , \qquad s = m , \tag{73}$$

$$K_{*s}^{*i} = \sum_{r \leqslant s} K_{rs}^{ii} + \sum_{j \neq i} \sum_{r \leqslant s} K_{rs}^{ji} , \qquad s > m . \tag{74}$$

If more detailed information is required, such as K_{rs}^{*i}, then this can be obtained by summing the appropriate elements of the full accounting matrix, **K**, all of which have now been calculated. *Steps 4-7* must be repeated until convergence is achieved.

This completes the description of the account-based model for the historical case. We should note that it is a straightforward matter to adapt it for projection. Although the birth rates b_{r*}^{*i} and b_{*u}^{*i} were calculated in *step 5*, they are not essential to the model in the historical case. It is useful to calculate them for a series of historical periods, however, so that trends can be analysed. In the projection case the various steps are modified as follows:

Step 1: Assemble data as before, but now assume rates d_{*s}^{*i} and b_{*u}^{*i} to be given rather than the death and birth totals. Assume that initial population is given and that migration rates can be projected, say using some model.

Step 2: K_{rs}^{ij}, $i \neq j$, is formed as before, but not $K_{r*}^{\beta(i)*}$.

Step 3: Initial values of at-risk populations can be obtained using known flows. Designate these as $\hat{K}_{*s}^{\mathrm{D}*i}$(init.) and $\hat{K}_{*u}^{\mathrm{B}*i}$(init.). Then initial values of total deaths and births can be obtained as:

$$K_{*s}^{*\delta(i)} = d_{*s}^{*i} \hat{K}_{*s}^{\mathrm{D}*i}(\text{init.}) , \tag{75}$$

and

$$K_{**}^{\beta(i)*}(K_{*u}^{*\mathrm{m}(i)}) = b_{*u}^{*i} \hat{K}_{*u}^{\mathrm{B}*i}(\text{init.}) . \tag{76}$$

Then $K_{r*}^{\beta(i)*}$(init.) can be found using equation (38).

Step 4: Calculate initial values for the major flows by using flow estimates obtained or calculated in *steps 1-3* in equations (40)-(50).

Step 5: Calculate the at-risk populations $\hat{K}_{*s}^{\mathrm{D}*i}$, $\hat{K}_{*s}^{\beta*i}$ using equations (52)-(54).

Step 6: Calculate total deaths and births from

$$K_{*s}^{*\delta(i)} = d_{*s}^{*i} \hat{K}_{*s}^{\mathrm{D}*i} , \tag{77}$$

$$K_{*u}^{\beta(i)*} = b_{*u}^{*i} \hat{K}_{*u}^{\mathrm{B}*i} , \tag{78}$$

and $K_{r*}^{\beta(i)*}$, using equation (38).

Step 7: We cannot now use equation (59) to calculate the unknown flows $K_{rs}^{i\delta(j)}$, $i \neq j$, since the rate d_{rs}^{*j} is unavailable. We therefore use the rate which is available, d_{*s}^{*j}, and calculate an at-risk population $\hat{K}_{*s}^{\mathrm{D}i\delta(j)}$ obtained by summing $\hat{K}_{rs}^{\mathrm{D}i\delta(j)}$

in equation (32) over r. We begin by restricting ourselves to the case $s > m$:

$$\hat{K}_{*s}^{\mathrm{D}i\delta(j)} = \sum_{r,u \geqslant s} {}_s^j\theta_{ru}^{\mathrm{D}ij} K_{ru}^{ij} + \sum_{r,u \geqslant s} {}_s^j\theta_{ru}^{\mathrm{D}i\delta(j)} K_{ru}^{i\delta(j)} , \qquad s > m . \tag{79}$$

We can substitute this into the equation

$$\frac{K_{*s}^{i\delta(j)}}{\hat{K}_{*s}^{\mathrm{D}i\delta(j)}} = d_{*s}^{*j} , \qquad s > m , \tag{80}$$

and rearrange to solve for $K_{*s}^{i\delta(j)}$ as follows:

$$K_{*s}^{i\delta(j)} = d_{*s}^{*j}\left(\sum_{r,u \geqslant s} {}_s^j\theta_{ru}^{\mathrm{D}ij} K_{ru}^{ij} + \sum_{r,u \geqslant s} {}_s^j\theta_{ru}^{\mathrm{D}i\delta(j)} K_{ru}^{i\delta(j)} \right) , \qquad s > m . \tag{81}$$

$K_{rs}^{i\delta(j)}$ can then be obtained using c-coefficients as

$$K_{rs}^{i\delta(j)} = c_{rs} K_{*s}^{i\delta(j)} , \qquad s > m . \tag{82}$$

As usual, equations (81) and (82) have to be solved iteratively for $K_{*s}^{i\delta(j)}$ as there are $K_{ru}^{i\delta(j)}$ terms on the right hand side of equation (81).

There is a similar problem in relation to the $K_{rs}^{\beta(i)\delta(j)}$ terms. We must tackle this problem simultaneously with the $K_{rs}^{i\delta(j)}$ problem for $s \leqslant m$. Equation (60) contains d_{0s}^{*j}, while what we have available is, again, d_{*s}^{*j}. d_{*s}^{*j}, for $s \leqslant m$, refers to the deaths $\sum_r K_{rs}^{\beta(i)\delta(j)} + \sum_r K_{rs}^{i\delta(j)}$. If we denote the corresponding at-risk population by $K_{*s}^{\mathrm{D}\beta(i)\delta(j)/i\delta(j)}$, it is given by

$$\begin{aligned}\hat{K}_{*s}^{\mathrm{D}\beta(i)\delta(j)/i\delta(j)} = {} & \sum_{u \geqslant s} {}_s^j\theta_{0u}^{\mathrm{D}\beta(i)j} K_{*u}^{\beta(i)j} + \sum_{u \geqslant s} {}_s^j\theta_{0u}^{\mathrm{D}\beta(i)\delta(j)} K_{*u}^{\beta(i)\delta(j)} + \sum_{r,u \geqslant s} {}_s^j\theta_{ru}^{\mathrm{D}ij} K_{ru}^{ij} \\ & + \sum_{r,u \geqslant s} {}_s^j\theta_{ru}^{\mathrm{D}i\delta(j)} K_{ru}^{i\delta(j)} , \qquad s \leqslant m .\end{aligned} \tag{83}$$

We can substitute this into

$$\frac{K_{*s}^{\beta(i)\delta(j)} + K_{*s}^{i\delta(j)}}{\hat{K}_{*s}^{\mathrm{D}\beta(i)\delta(j)/i\delta(j)}} = d_{*s}^{*j} , \qquad s \leqslant m , \tag{84}$$

and solve for $K_{*s}^{\beta(i)\delta(j)} + K_{*s}^{i\delta(j)}$ to give

$$\begin{aligned}K_{*s}^{\beta(i)\delta(j)} + K_{*s}^{i\delta(j)} = d_{*s}^{*j} \Bigg(& \sum_{u \geqslant s} {}_s^j\theta_{0u}^{\mathrm{D}\beta(i)j} K_{*u}^{\beta(i)j} + \sum_{u \geqslant s} {}_s^j\theta_{0u}^{\mathrm{D}\beta(i)\delta(j)} K_{*u}^{\beta(i)\delta(j)} + \sum_{r,u \geqslant s} {}_s^j\theta_{ru}^{\mathrm{D}ij} K_{ru}^{ij} \\ & + \sum_{r,u \geqslant s} {}_s^j\theta_{ru}^{\mathrm{D}i\delta(j)} K_{ru}^{i\delta(j)} \Bigg) , \qquad s \leqslant m .\end{aligned} \tag{85}$$

The equation (85) must be solved iteratively as usual. We can then subdivide this $K_{*s}^{\beta(i)\delta(j)} + K_{*s}^{i\delta(j)}$ term using c- and e-coefficients as we did for $K_{*s}^{\beta(i)\delta(i)} + K_{*s}^{i\delta(i)}$ in the main accounting equations:

$$K_{*s}^{\beta(i)\delta(j)} = c_{0s}(K_{*s}^{\beta(i)\delta(j)} + K_{*s}^{i\delta(j)}) , \tag{86}$$

$$K_{*s}^{i\delta(j)} = (1 - c_{0s})(K_{*s}^{\beta(i)\delta(j)} + K_{*s}^{i\delta(j)}) , \tag{87}$$

$$K_{rs}^{\beta(i)\delta(j)} = e_{rs} K_{*s}^{\beta(i)\delta(j)} , \tag{88}$$

$$K_{rs}^{i\delta(j)} = c_{rs} K_{*s}^{i\delta(j)} . \tag{89}$$

Step 8: The accounting equations (61)–(74) can now be solved in the usual way.

5 The transition-rate form of the model

Any account-based model can be presented in terms of transition rates. We can show this as follows: let

$$K^{\hat{}*} = \begin{bmatrix} K_{1*}^{1*} \\ K_{2*}^{1*} \\ \cdot \\ K_{1*}^{2*} \\ \cdot \\ K_{\lambda*}^{\beta(1)*} \\ \cdot \\ K_{\mu*}^{\beta(1)*} \\ \cdot \\ K_{\lambda*}^{\beta(2)*} \\ \cdot \end{bmatrix}, \quad \text{and} \quad K^{*\hat{}} = \begin{bmatrix} K_{*1}^{*1} \\ K_{*2}^{*1} \\ \cdot \\ K_{*1}^{*2} \\ \cdot \\ K_{*1}^{*\delta(1)} \\ K_{*2}^{*\delta(1)} \\ \cdot \\ \cdot \\ K_{*1}^{*\delta(2)} \\ \cdot \end{bmatrix} \tag{90}$$

denote vectors which are of course the row and column sums of the accounts matrix **K**. Let **G**$'$ (where the prime denotes transposition) be the matrix of transition rates formed by dividing each element of **K** by the corresponding row sum. Then it can easily be checked that the following relationship holds identically:

$$K^{*\hat{}} = \mathbf{G}K^{\hat{}*} . \tag{91}$$

This matrix can also be used to generate the accounts matrix as

$$K' = \mathbf{G}\widetilde{\mathbf{K}}^{\hat{}*} , \tag{92}$$

where $\widetilde{\mathbf{K}}^{\hat{}*}$ is the matrix whose diagonal elements are the elements of $K^{\hat{}*}$, and with zeros elsewhere. Generally speaking, equation (91), rather than (92), will be called the transition-rate model. **G** is an operator which transforms the 'old population' vector, $K^{\hat{}*}$, into a 'new population' vector, $K^{*\hat{}}$. Of course, equation (91) is simply an algebraic identity. It only becomes a *model* equation if the rates, the elements of **G**, can be supplied by some independent means. In terms of the main purposes of this paper, there is little to gain from this formulation. However, it does provide a useful basis for comparison between our own account-based model and others. We shall illustrate this argument by indicating how our own model can be compared with that of Rogers (1966, 1968). We begin by showing how the transition-rate model given by equation (91) can be manipulated to facilitate the comparison.

The rates associated with the $\mathbf{K}^{ii}$, $\mathbf{K}^{ij}$, $i \neq j$, $\mathbf{K}^{i\delta(i)}$, and $\mathbf{K}^{i\delta(j)}$ submatrices in **K** are all obtained by dividing by base year populations, and are recognisable as kinds of survival, migration and survival, death, and migration and death rates respectively. Those associated with all kinds of births are obtained by dividing by the corresponding total number of births. They are not obviously recognisable as birth rates because of these denominators. We can amend this situation as follows. Simultaneously replace elements $K_{r*}^{\beta(i)*}$ in $K^{\hat{}*}$, and in the denominators of corresponding elements of **G**, by K_{r*}^{i*}, the appropriate base year population[2]. $K^{\hat{}*}$ now has the final NR rows identical to the first NR, so we can contract equation (91) by forming a matrix $\hat{\mathbf{G}}$ from **G**. This is obtained by adding the $(NR+j)$th column of **G** to the jth to form the $2NR \times NR$ matrix $\hat{\mathbf{G}}$. If we do not wish to record deaths explicitly, then we can contract still further by deleting the final NR rows in the vector $K^{*\hat{}}$, and the final NR rows in $\hat{\mathbf{G}}$ to form $\hat{\hat{\mathbf{G}}}$. With the new and old population vectors modified in this way, the transition-rate model then

[2] For some purposes it may be appropriate to use $\hat{K}_{r*}^{\mathrm{B}*i}$ rather than K_{r*}^{i*} (see Rees and Wilson, 1974, chapter 12), but here we use K_{r*}^{i*} to facilitate later comparisons.

becomes

$$K^{*\wedge} = \hat{\mathbf{G}}K^{\wedge *} . \tag{93}$$

It can easily be checked that the model in this form is directly comparable with that of Rogers. We note that Rogers always makes the simplifying assumptions of equal age-group intervals, each equal to the projection period T, and so the reader who wishes to make the detailed comparison for himself should apply this assumption to any detailed specification of equation (93). Here we only present the bare bones of the result of such a comparison, the details being given elsewhere (Rees and Wilson, 1974). First, the nonzero terms in the matrix **G** are in the same place in the transition-rate model, given by equation (93), and in Rogers' model—as birth, survival, and migration terms—with one exception: terms such as $K_{r1}^{\beta(i)j}/K_{1*}^{i*}$ appear in the transition-rate model to represent migrating infants. Such rates could easily be added to the Rogers' model. Second, and more important, the transition-rate model provides a proper definition of the rates to be used in such a model. For example, survival rates are of the form $K_{rr+1}^{ii}/K_{r*}^{i*}$, birth rates of the form $K_{r1}^{\beta(i)}/K_{r*}^{i*}$, migration rates of the form $K_{rr+1}^{ij}/K_{r*}^{i*}$, and death rates of the form $K_{rr}^{i\delta(i)}/K_{r*}^{i*}$ or $K_{rr+1}^{i\delta(i)}/K_{r*}^{i*}$. It is unlikely that most users of the Rogers' model define their rates in this way, particularly because of data availability. It is tempting, for example, to use $K_{r1}^{\beta(i)*}$ as the numerator in the birth-rate definition simply because it is available from data. Many errors, some slight, some not so slight, are commonly introduced in this way.

These problems can be resolved by use of the account-based model of section 4. Of the numerators listed for the rates above, only K_{rr+1}^{ij} is directly available from data, but the others can be obtained using the account-based model. (Though the rates *are* transition rates, and they will not be the same as those calculated in the account-based model using at-risk populations.) Of course, there is little to be gained from this procedure—if transition rates were calculated in this way, the final predictions of the transition-rate model would merely replicate those which had already been obtained from the account-based model.

There is a third important point which relates to projection. In projecting with the account-based model, we use birth and death rates which can be obtained from an analysis of historical values of such rates. Thus, the rates used connect directly to available data in historical time series, and are simpler in structure, and more likely to be represented by stable trends for projection, than the whole array of rates needed for projection with the transition-rate model.

6 Sex disaggregation

So far we have concentrated on age-disaggregation and assumed that we have had a single-sex population which can conveniently be assumed to consist of females. We now show how to analyse a two-sex population.

For most purposes it suffices to define a full accounting matrix for each sex, say as $\mathbf{K}^{♀}$ and $\mathbf{K}^{♂}$, to assume that there is no sex-change, and thus to have an overall accounting matrix, **K**, given by

$$\mathbf{K} = \left[\begin{array}{c|c} \mathbf{K}^{♀} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{K}^{♂} \end{array}\right], \tag{94}$$

where **0** is a matrix of 0's of the same dimensions as $\mathbf{K}^{♀}$ and $\mathbf{K}^{♂}$ (that is, $2NR \times 2NR$).

The accounting and other equations can be extended in a straightforward way by adding ♂ (male) or ♀ (female) superscripts as appropriate. Care must be taken in relation to one point, however: all equations for at-risk populations *for births* involve female populations only.

Finally, in building a sex-disaggregated transition-rate model, one further adjustment is needed. Any accounting matrix K can be divided into two $NR \times 2NR$ submatrices E (standing for *e*xistence) and B (standing for *b*irth):

$$\mathbf{K} = \left[\begin{array}{c} \mathbf{E} \\ \hline \mathbf{B} \end{array}\right] . \tag{95}$$

Then this would lead to a two-sex population being represented [compare with equation (94)] as

$$\mathbf{K} = \left[\begin{array}{c|c} \begin{matrix}\mathbf{E}^{♀} \\ \mathbf{B}^{♀}\end{matrix} & 0 \\ \hline 0 & \begin{matrix}\mathbf{E}^{♂} \\ \mathbf{B}^{♂}\end{matrix} \end{array}\right] . \tag{96}$$

For the purposes of building the transition model, and in particular for carrying out the series of operations which lead to the Rogers' model, it is more convenient to change the position of the $\mathbf{B}^{♂}$ matrix as follows:

$$\mathbf{K} = \left[\begin{array}{c|c} \begin{matrix}\mathbf{E}^{♀} \\ \mathbf{B}^{♀}\end{matrix} & \begin{matrix}0 \\ \mathbf{B}^{♂}\end{matrix} \\ \hline 0 & \begin{matrix}\mathbf{E}^{♂} \\ 0\end{matrix} \end{array}\right] . \tag{97}$$

The process described earlier to construct a transition-rate model and to compress the result into a Rogers' type model, can then be applied to this matrix.

7 Concluding comments

Our aim in this paper has been to explain in broad outline the main concepts associated with an age-sex disaggregated account-based spatial demographic model. Because of space limitations much detail, inevitably, has had to be omitted. We have mainly concentrated on the general case. In practice much of our empirical work has been focused on the simple case. There are two further forthcoming papers which the reader may like to pursue to fill in some of the gaps. First, we note that much of the design of our model has been conditioned by data availability. These issues are explored much more fully in Rees and Wilson (1973b). Second, the argument can probably be classified in relation to empirical work. A detailed description of the building of a 'simple case' disaggregated model for West Yorkshire is presented in Smith and Rees (1973). Full details of all aspects of spatial demographic analysis from the viewpoint of the account-based model will be presented in a forthcoming book (Rees and Wilson, 1974).

References

Rees, P. H., 1973, "A revised notation for spatial demographic accounts and models", *Environment and Planning,* **5**, 147-155.

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

Rees, P. H., Wilson, A. G., 1973b, "Population statistics and spatial demographic accounts", Working Paper 53, Department of Geography, University of Leeds.

Rees, P. H., Wilson, A. G., 1974, *Spatial Demographic Analysis*, mimeo.

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

Rogers, A., 1968, *Matrix Analysis of Inter-regional Population Growth and Distribution* (University of California Press, Berkeley).

Smith, A. P., Rees, P. H., 1973, "Methods for constructing spatial demographic accounts for a British region", Working Paper 59, Department of Geography, University of Leeds.

Wilson, A. G., 1972, "Multiregional models of population structure and some implications for a dynamic-residential location model", in *London Papers in Regional Science 3: Patterns and Processes in Urban and Regional Systems,* Ed. A. G. Wilson (Pion, London), pp.217-242.

© 1974 a Pion publication printed in Great Britain