Working Paper 295/

Computer Manual 13

ACCOUNTS BASED MODELS FOR MULTIREGIONAL POPULATION ANALYSIS: METHODS, PROGRAM AND USERS' MANUAL

P.H. Rees

School of Geography
University of Leeds
Leeds, LS2 9JT
U.K.

January, 1981

CONTENTS

* Not included as such with the program. Instructions for their listing are supplied.

Preface

This manual describes the accounts based model program ABM) as developed to December 31st, 1980. All su .routines have been tested out with the most likely settings of the analysis parameters, and should work efficiently for virtually all users.

However, not all parameter combinations have been tested out,* and some will result in failure because the population model specified is erroneous. Error checking routines will be added as these erroneous parameter combinations detected.

Users having difficulty with the program should contact me.

Philip Rees

31st December, 1980.

* The reason for this is the very large number of parameter value combinations possible (exceeding 1.3×10^8)!

Acknowledgements

The author is very grateful to the University of Leeds on two counts:

Firstly, the School of Geography and the Registrar permitted him one term's study leave in which to concentrate on the development of the program. Long and complicated programs cannot be developed without such sustained and undistracted effort.

Secondly, the University Computing Service have provided computer users with a machine, the Amdahl VM470, and a campus network of terminals and printers, that perform so well that the critical bottleneck of program writing, development and testing is no longer machine availability or machine time but the capacity of the programmer for logical thought. This is how it should be.

Finally, the author owes a particular debt to Frans Willekens and Andrei Rogers whose "Spatial Population Analysis: Methods and Computer Programs (Research Report RR-78-18)" provides the model of manual writing upon which the current work is based. The ABM program complements the analysis of the Willekens and Rogers programs.

1. BACKGROUND

1.1 A bit of history

The ABM (accounts based model) program replaces an earlier program DAME (demographic accounts model estimator) (Plessis-Fraissard and Rees, 1977a, 1977b, 1978) that estimated age-sex disaggregated population accounts and incorporates the flexible projection capacity of the MULTIPOP (multiregional population) programs (Jenkins and Rees, 1977, 1978) for constructing aggregate population accounts. Most of the disadvantages inherent in earlier programs have been overcome and in addition the program has been designed to work with single year of age - single year of time data in time for use with the new migration information to be generated in the 1981 U.K. Census.

1.2 Features of the program

1.2.1 Accounts

The program ABM (accounts based model) makes estimates of population accounts for multiregional systems for past or future periods of time. The population accounts can be prepared for the whole or aggregate populations or sex disaggregated population or age disaggregated populations or age and sex disaggregated populations. Population accounts contain data on the populations of the regions being studied and the components that change those populations from both a multi-regional and single region point of view.

1.2.2 Inputs

The inputs to the program consist of selected information on the components that make up the accounts (see section 4). A wide variety of different "data situations" and associated models can be handled by the program, which should make it more widely applicable. The standard inputs to the accounts based model consist in part of

initial populations
surviving internal migrants
surviving immigrants
surviving emigrants
deaths

However, information on surviving emigrants may not be available so that alternative inputs can be used.

initial populations
surviving internal migrants
surviving immigrants
deaths
final population.

The population and migration information required should be available from successive censuses (under certain conditions). Or, alternateively, the information may be in the form of population movements rather than transitions.

mid-period population
internal migrations
immigrations
emigrations
deaths.

This "data situation" can be handled by the program as long as a set of empirical ratios (of movements to transitions) can be specified or if the assumptions of a model effecting that conversion are accepted.

1.2.3 The bases of projections

In carrying out projections of the population rates may be substituted for flows in estimating the components of change or the results of other models (e.g. migration models) could be input.

1.2.4 Mimicry of simpler models

By judicious selection of the inputs to the program simpler models such as the single region cohort survival with net migration flows or rates can be mimicked so that users can trace the effect of adopting more sophisticated models independent of their assumptions about population behaviour.

1.2.5 Flexibility : the key feature

The overall flavour of the program is therefore flexibility : ABM enables users to tailor the accounts based model they build to their data resources and information needs.

1.2.6 Reporting of the analysis results

This flexibility is carried over into the reporting of results by the program. A variety of different levels of report are provided: the most detailed levels (e.g. information on all iterations of the accounts based model or of a constraints nature) are intended only as an aid to error detection; the least detailed level provides information only on population numbers in the regions at points in time.

1.2.7 Estimation of data inputs

In order to retain this flexible nature, a good deal of data estimation has been left outside the program. The user must design his or her own methods for converting the information available to that required by ABM, which will, of course, be particular to the regional system being studied. A set of programs are being developed in the context of British data that carries out such estimations which the user may wish to use or base his or her own programs on

2. THEORY

2.1 Population accounts

The concepts of population accounts have been developed by the Cambridge economist Richard Stone (Stone, 1971, 1975) and by Rees and Wilson (1977). The reader is referred to the latter book and subsequent papers (Rees, 1979, 1980a, 1980b) for a full exposition of the full theory underpinning population accounts. Here a simplified description is provided of population accounts for cohorts. Since the program constitutes a model building as well as a data processing system, the user should have a reasonable grasp of the principles involved before utilizing ABM (particularly, Part 2 of Rees and Wilson, 1977).

2.1.1 Notation

The accounts matrix is a two dimensional array of population numbers classified into cells according to their initial states in a time interval (the row classes) and their final states in a time interval (the column classes), with other states such as sex or, in this paper, age cohort, constant over the time interval.

General notation for population

K : a count of people

Initial life states

v : general label for initial life-state of which there are two:-

b : birth in a time interval

e : existence at the start of a time interval

Final life states

w : general label for final life-state of which there are two:-

s : survival at the end of a time interval

d : death in a time interval

Constant states

a : age cohort label. There are a = 1, ..., A cohorts. The concept is depicted in Figure 2.1. Persons remain in the age cohort for a time interval and are transferred to the next, assuming survival, at the start of the next time interval. They remain members of the same cohort of persons born in the same set of birth dates throughout.

This age cohort definition represents a return to that suggested by Stone (1971).

x : sex label with values m for males, f for females and p for persons. Sex labels are generally not made explicit as most accounting equations refer to a single sex.

Other constant (or fairly constant) states could also be specified: race or ethnic status for example, although over several periods interaction will take place.

Time labels

t : start point of time interval

T : length of time interval

t+T : end point of time interval

These labels are generally not made explicit as most accounting equations refer to a single period t to t + T.

Initial and final states of interest

In this paper we use:-

i : initial region label for region at start of time interval

j : final region label for region at end of time interval

There are i,j = 1, ..., N internal regions of interest and R is the label used for the external region or 'rest of the world'.

People can move freely among these region states. Note that we could use the labels i and j to refer to other kinds of states amongst which people can move more or less freely such as:

illness states (cf Chiang, 1968)

income groups

occupations

marital statuses (although some transitions may not be allowed)

household types

labour force status (employed, unemployed, inactive, retired and so on)

educational attainment states (though the most interesting transitions occur before age 25 with education becoming a relatively constant classification thereafter)

Thus, although accounts are defined and discussed in this paper as referring to regional age-cohort disaggregated populations, these are only one of the many sorts of population that could be described using accounting theory.

2.1.2 The accounts variables

At the most general level, two accounts variables can be defined, survivors and non-survivors.

$K_a^{v(i)s(j)}$: persons of age cohort either born in region i or in existence in region i who survive in region j at the end of the period (survivors)

$K_a^{v(i)d(j)}$: persons of age cohort a either born in region i or in existence in region i who die in region j before the end of the time interval (non-survivors)

If these variables are summed over final and initial states we obtain the row and column sums of the accounts:

$K_a^{v(i)*(*)}$: row sums of the accounts matrix

$K_a^{*(*)s(j)}$: column sums for survivors of the accounts matrix

$K_a^{*(*)d(j)}$: column sums for non-survivors of the accounts matrix

2.1.3 The accounts matrix and table

These variables can be arranged in a matrix with 2n columns and n rows, where n is the number of regions being studied (Figure 2.2). When row sums and columns are added to the accounts matrix, an accounts table is formed.

The accounts table becomes of real interest only when we define the set of regions being studied to include an external zone that receives or sends all the population flows which the internal regions send to 'the rest of the world'. Then we can interpret the row sums for the internal regions as the initial populations of the regions or the births totals in those regions.

Figure 2.3 shows a four region accounts table in which the last region is the external zone that closes the system being studied together with a numerical example. In these accounts the population transitions within the rest of the world are ignored and set to zero. The row sums of these accounts can be interpreted as initial populations of the regions (in the example: 125605, 1409274, 3058068) except for the last (113704) in which the total refers to the sum of all surviving and non-surviving immigrants to the internal system. The column sums for survivors are the final populations of the regions in the time interval (in the example: 135051, 1356147, 3016295), except for the last which refers to the total of surviving emigrants from all the internal regions (177998). The non-survivors totals refer to deaths recorded in the internal regions (530,5708,14525) except for the rest of the world total which refers to non-surviving emigrants (397).

2.2 Components of the accounts

In order to make estimates of the numbers of persons experiencing the transitions laid out in an accounts table, we need to identify those items of information which can be acquired from official demo-

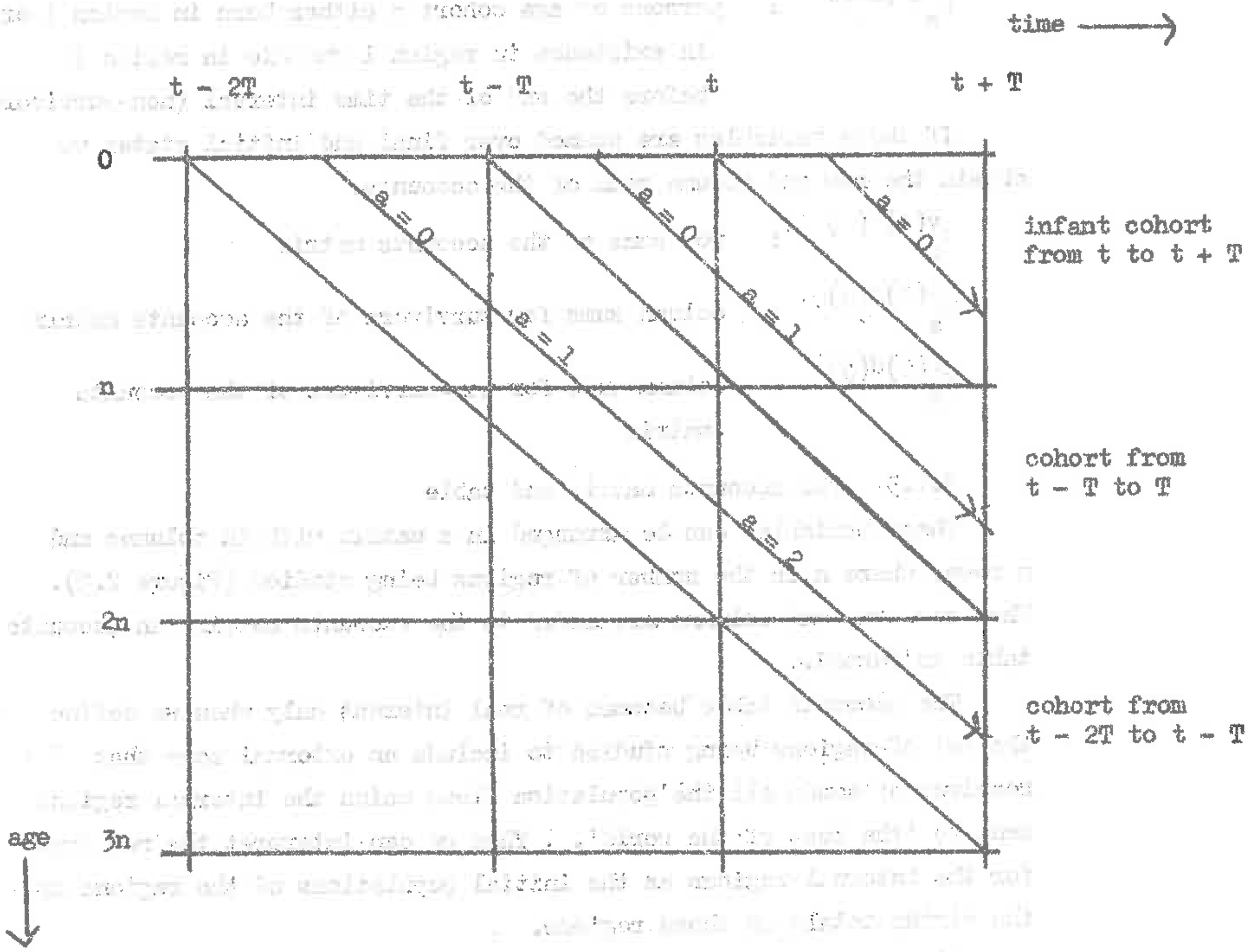


Figure 2.1 The age cohort concept illustrated

Initial states \ Final states		Region of survival (s) at time t + T						Region of death (d), t to t + T						Totals
		1	2	...	j	...	n	1	2	...	j	...	n	*
Initial state (v): existence (e) or birth Region of origin	1	$K_a^{v(1)s(1)}$	$K_a^{v(1)s(2)}$	...	$K_a^{v(1)s(j)}$	...	$K_a^{v(1)s(n)}$	$K_a^{v(1)d(1)}$	$K_a^{v(1)d(2)}$	...	$K_a^{v(1)d(j)}$	...	$K_a^{v(1)d(n)}$	$K_a^{v(1)*(*)}$
	2	$K_a^{v(2)s(1)}$	$K_a^{v(2)s(2)}$	...	$K_a^{v(2)s(j)}$	...	$K_a^{v(2)s(n)}$	$K_a^{v(2)d(1)}$	$K_a^{v(2)d(2)}$	...	$K_a^{v(2)d(2)}$	...	$K_a^{v(2)d(n)}$	$K_a^{v(2)*(*)}$
	⋮	⋮	⋮		⋮		⋮	⋮	⋮		⋮		⋮	⋮
	i	$K_a^{v(i)s(1)}$	$K_a^{v(i)s(2)}$	...	$K_a^{v(i)s(j)}$	...	$K_a^{v(i)s(n)}$	$K_a^{v(i)d(1)}$	$K_a^{v(i)d(2)}$		$K_a^{v(i)d(j)}$	...	$K_a^{v(i)d(n)}$	$K_a^{v(i)*(*)}$
	⋮	⋮	⋮		⋮		⋮	⋮	⋮		⋮		⋮	⋮
	n	$K_a^{v(n)s(1)}$	$K_a^{v(n)s(2)}$	...	$K_a^{v(n)s(j)}$	...	$K_a^{v(n)s(n)}$	$K_a^{v(n)d(1)}$	$K_a^{v(n)d(2)}$	...	$K_a^{v(n)d(j)}$	...	$K_a^{v(n)d(n)}$	$K_a^{v(n)*(*)}$
totals	*	$K_a^{v(x)s(1)}$	$K_a^{v(x)s(2)}$	...	$K_a^{v(x)s(j)}$	...	$K_a^{v(*)s(n)}$	$K_a^{v(x)d(1)}$	$K_a^{v(x)d(2)}$		$K_a^{v(*)d(j)}$		$K_a^{v(*)d(n)}$	$K_a^{v(*)*(*)}$

Figure 2.2 <u>The structure of a population accounts table for age cohort r</u>

graphic sources (or estimated fairly reliably from such sources) and which will need to be estimated by means of 'model' equations (guesses as to their likely size) or by means of accounting equations (that is, as residuals in the row or column in which they appear, all other items being known).

To do this the variables are classified into further component parts. These components are shown in Figure 2.4 in words using the four region accounts of Figure 2.3 as the framework.

Some twelve components are identified, all of which must be estimated to complete any accounts table, though one, the populations at risk, do not appear as such in the table. Each component can be estimated in one of a number of different ways, and these choices are discussed, component by component, in the third section of the paper. For example, non-surviving internal migrants could be estimated using one of three hypotheses: that migrants die at the rate observed in their region of destination; that migrants die at the rate observed in their region of origin; or that migrants die at a rate intermediate between that of their origin and their destination regions.

2.3 General features of the accounts based model

The combination of choices for estimating items in each component constitutes an accounts based model. Not all choice combinations are allowed but there are a great many resultant accounts based models. The detailed equations involved are not given here, but may be assembled from those described in section 3. All will have the following features, however.

(1) The populations of the regions are input externally. The initial populations may be derived from population census data or population estimate series or they may simply be the final populations of the previous modelling period (if forecasting is being carried out). The final populations may be derived from population census data or population estimate series or they may simply be the initial populations of the next modelling period (if backcasting is being carried out). Either initial populations or final populations or both may be input.

(2) Surviving internal migrant flows will be either derived from census data or registration sources and input as flows or they will be estimated by inputting migration rates and multiplying these by the initial population serving as the population at risk. Or they may be estimated by a spatial interaction model and then input.

Figure 2.3 The structure of an accounts matrix for three internal and one external region, together with a numerical example for the 0-4 cohort in 1966-71

3A

Initial state \ Final state		Region of survival (s) at time t + T				Region of death (d), t to t + T				Totals	Interpretation of totals
		1	2	3	4	1	2	3	4		
Existence in region of origin	East Anglia 1	$K_a^{e(1)s(1)}$	$K_a^{e(1)s(2)}$	$K_a^{e(1)s(3)}$	$K_a^{e(1)s(4)}$	$K_a^{e(1)d(1)}$	$K_a^{e(1)d(2)}$	$K_a^{e(1)d(3)}$	$K_a^{e(1)d(4)}$	$K_a^{e(1)*(*)}$	Initial populations
	South East 2	$K_a^{e(2)s(1)}$	$K_a^{e(2)s(2)}$	$K_a^{e(2)s(3)}$	$K_a^{e(2)s(4)}$	$K_a^{e(2)d(1)}$	$K_a^{e(2)d(2)}$	$K_a^{e(2)d(2)}$	$K_a^{e(2)d(4)}$	$K_a^{e(2)*(*)}$	
	Rest of Britain 3	$K_a^{e(3)s(1)}$	$K_a^{e(3)s(2)}$	$K_a^{e(3)s(3)}$	$K_a^{e(3)s(4)}$	$K_a^{e(3)d(1)}$	$K_a^{e(3)d(2)}$	$K_a^{e(3)d(3)}$	$K_a^{e(3)d(4)}$	$K_a^{e(3)*(*)}$	
	Rest of World 4	$K_a^{e(4)s(1)}$	$K_a^{e(4)s(2)}$	$K_a^{e(4)s(3)}$	0	$K_a^{e(4)d(1)}$	$K_a^{3(4)d(2)}$	$K_a^{e(4)d(3)}$	0	$K_a^{e(4)*(*)}$	Immigrants
Totals		$K_a^{e(*)s(1)}$	$K_a^{e(*)s(2)}$	$K_a^{e(*)s(3)}$	$K_a^{e(*)s(4)}$	$K_a^{e(*)d(1)}$	$K_a^{e(*)d(2)}$	$K_a^{e(*)d(3)}$	$K_a^{e(*)d(4)}$	$K_a^{e(*)*(*)}$	
Interpretation of totals		Final populations			Surviving emigrants		Deaths	Non-surviving emigrants	Grand total		

THE ACCOUNTS MATRIX PERIOD = 1 AGE = 1 SEX = 1 ITERATION = 3

	EAST ANG 1	SOUTH EA 2	REST BRT 3	REST WLD 4	EAST ANG 5	SOUTH EA 6	REST BRT 7	REST WLD 8	TOTALS 9
EAST ANG	108108.	5739.	6880.	4366.	475.	12.	16.	9.	125605.
SOUTH EA	11171.	1249289.	58934.	84055.	23.	5487.	141.	174.	1409274.
REST BRT	8869.	52833.	2892219.	89577.	18.	109.	14228.	214.	3058068.
REST WLD	6900.	48268.	58283.	0.	14.	[illegible]	140.	0.	113704.
TOTALS	135048.	1356129.	3016316.	177998.	530.	5738.	14525.	397.	4706551.

Figure 2.4 Components of the accounts

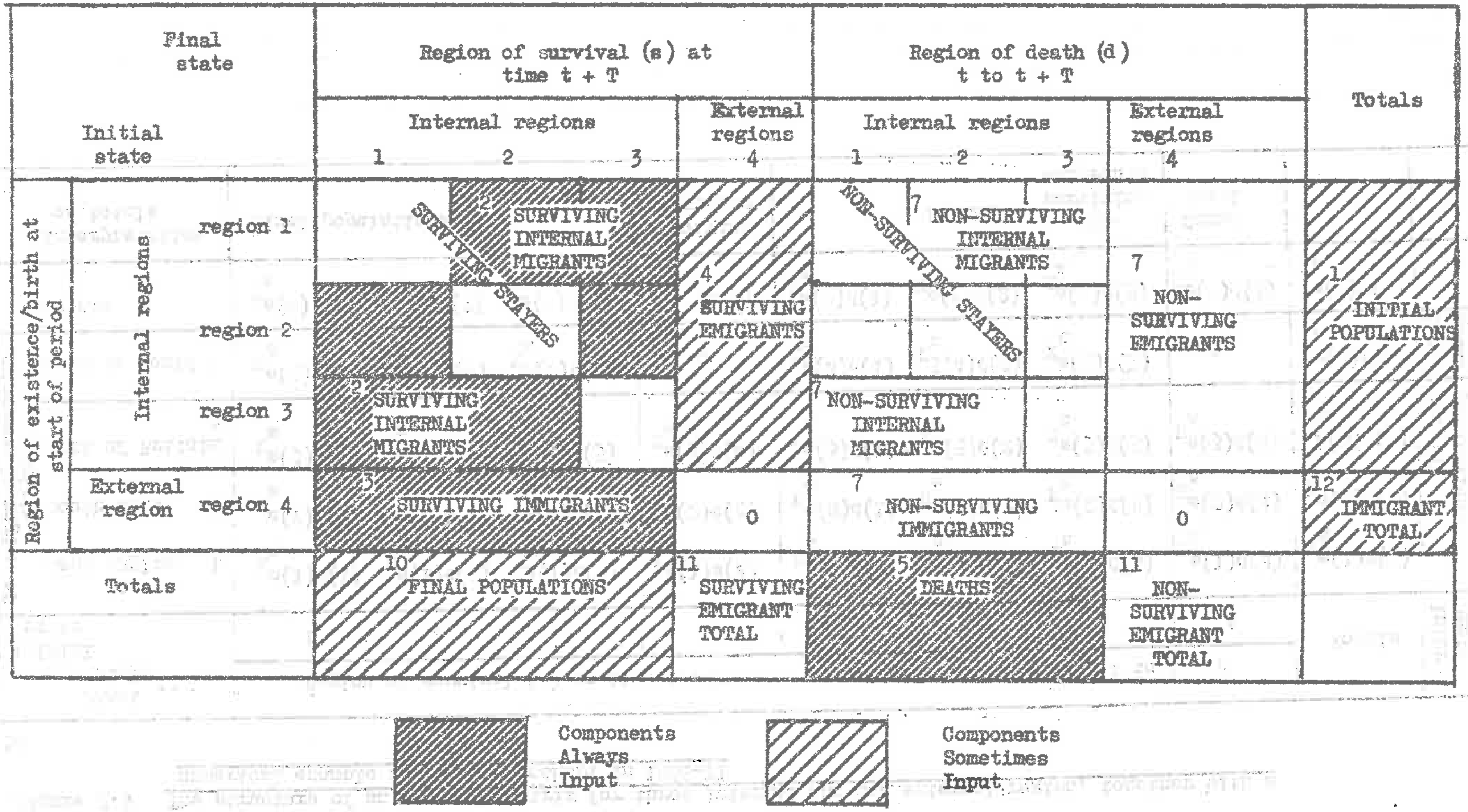
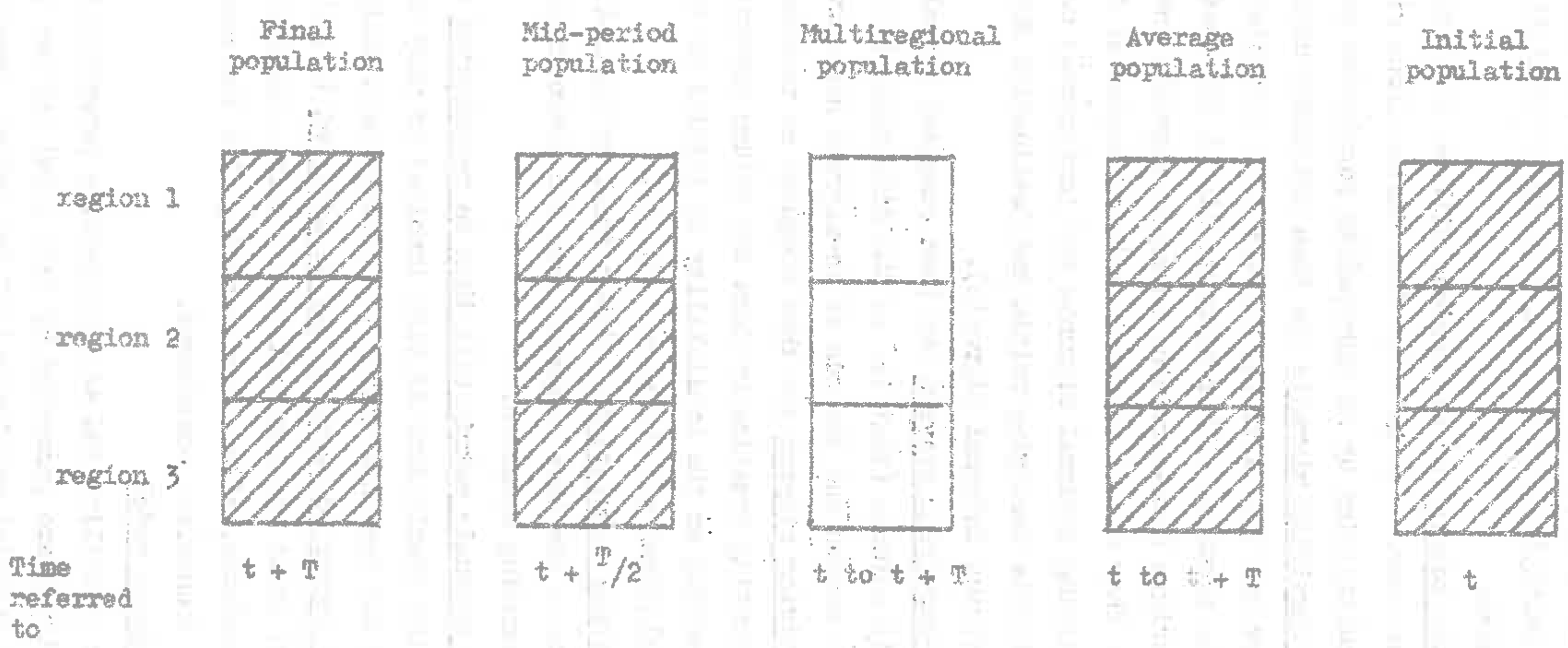


Figure 2.4 (continued)

6

POPULATIONS AT RISK



(3) Surviving external migrant flows (immigrants and emigrants) will be either input from census or registration or survey data sources. If emigrant estimates are unavailable, they may be derived by inputting both initial and final populations and immigrant flows. Rates may be substituted for flow inputs and multiplied by one of a variety of populations at risk instead.

(4) Deaths totals for the regions will be input in either flows or rates form. If in rates form deaths flows will be estimated by multiplying by an appropriate population at risk.

(5) Non-surviving migrants are then estimated using regional death rates and the surviving migrant flow estimates in model equations assuming either migrants die at the death rate of region of origin (non-surviving internal migrants and emigrants), or that they die at the death rate of region of destination (non-surviving internal migrants and immigrants) or that they die at a rate averaging origin and destination mortality (non-surviving internal migrants).

(6) Non-surviving stayers are then computed normally as residuals in their columns (that is, as the difference between total deaths in a region and the sum of in-migrant non-survivors).

(7) Surviving stayers are then computed as residuals from the row accounting equation or from the column accounting equation depending on the availability of initial or final regional populations and the users' choice.

(8) Regional populations, either final or initial, which remain unknown are then computed from the appropriate column or row accounting equation.

(9) The accounts matrix will then be reestimated using revised values of the populations at risk involved as many times as are required for the accounts matrix to converge on a stable solution.

(10) The converged accounts may be adjusted to new or unused information about the marginal totals using standard balancing factor procedures, after first ensuring that the marginal constraints are themselves consistent.

2.4 A births model

In order to utilize ABM as a projection program a births model is required. Births are generated by this model and are aggregated to represent the 'initial population' of the infant cohort. Accounts for the infant cohort are built in very much the same fashion as accounts for existing age cohorts except that several different options for some components are available and that the 'fraction of a period exposed to risk' factor in the population at risk equations is one half rather than one.

The births model either inputs births disaggregated by age cohort from data sources or inputs age cohort specific birth rates which are multiplied by the population at risk specified by the user. The age cohort classified births are summed before being used as the row totals (for internal regions) in the infant accounts.

2.5 Sex disaggregation

Births are disaggregated into two sexes by application of sex proportions (circa 52% for males and 48% for females in developed country populations) if required. The population at risk in the birth model may be specified to be the female population (female dominant model) or male population (male dominant model) or the two sexes may be treated separately or together.

Otherwise in the program there is no difference in the treatment of the sexes: separate information is input for each and separate accounts produced. These accounts may be aggregated into accounts for persons if required.

2.6 Data preparation for cohort accounts

The accounts based model for an age disaggregated population is considerably simplified if 'age cohorts' rather than 'age groups' are adopted as the units of age classification (Rees, 1978, 1980). The model equations will have approximately the form of those for the age aggregated model (Rees and Wilson, 1977, Part 2) rather than the more complex form of the age disaggregated model (Rees and Wilson, 1977, Part 3).

To take advantage of such simplification all input data must be prepared in 'age cohort' form. Since some demographic data are collected in age group form the problem of transformation to age cohort form requires solution. The process of data transformation has been left outside the ABM program because the exact requirements and techniques will differ from application to application. A couple of examples are included here to show that it is a relatively simple process.

Deaths data, for example, could be deconsolidated from single classification by age group to double classification by age group and age cohort form thus:

$$K_{aa}^{*(*)}\,d\,(j) = c_{aa}\;K_{*a}^{*(*)}\,d\,(j) \qquad (2.1)$$

$$K_{aa+1}^{*(*)}\,d\,(j) = c_{aa+1}\;K_{*a+1}^{*(*)}\,d\,(j) \qquad (2.2)$$

and then reaggregated to age cohort form

$$K_{a*}^{*(*)} d (j) = K_{aa}^{*(*)} d (j) + K_{aa+1}^{*(*)} d (j) \qquad (2.3)$$

The coefficients c_{aa} and c_{cc+1} could be derived from more detailed national data or from geometric considerations derived from the Lexis diagram (that is, c_{aa} and c_{aa+1} could both be set to 0.5). Or a more complex estimation model could be constructed (cf Rees, Smith and King, 1977).

2.7 Modes of use of the accounts based model

Initially, the accounts based model was designed either as an historical model in which all inputs were of known population stock or flow quantities or as a projection model in which all components besides initial populations were input as forecast rates. However, a much wider set of modes of use can be envisaged with the current accounts based model. These are set out in Figure 2.5.

If reasonable estimates of the numbers of people in each flow component input are available for a past time period, the resulting accounts are known as historical (the first row of the table in Figure 2.5). If no constraints (external estimates for the accounts table marginals)are applied the accounts are unconstrained. If the final population stocks are the ones available then the backcast model must be used (see the example in section 5). If both initial and final populations are known then constrained historical accounts will be prepared.

For one or more time intervals between the present and the period for which all input data are available, only some of input components may be known. In this situation estimated flows or rates may have to be input and the resulting accounts called 'estimated'. If the time interval being studied is wholly in the future then the accounts produced are referred to as projected as the inputs will all be forecast rather than known. In certain circumstances the analyst may wish to fix the population stocks, and the accounts based model may be used to estimate the relevant matrices of change that would match the target populations. All projected versions of the accounts based model can be used with fixed or variable inputs of rates or flows.

The detailed options and equations available for each component of the accounts are now specified in the next section of the paper.

Time Horizon	Knowledge of Flow Components	State of knowledge of stocks components		
		Initial population stocks only	Final population stocks only	Constraints (inc both initial and final population stocks)
Past	All flow data inputs are known	historical unconstrained	historical unconstrained backcast	historical constrained
	Only some flow data inputs are known	estimated unconstrained	estimated unconstrained backcast	estimated constrained
Present → Future	All flow data inputs are projected	projected un-constrained (forecast)	projected un-constrained backcast	projected constrained (target)

Figure 2.5 <u>Modes of use of the accounts based model</u>

3. THE BUILDING BLOCKS OF ABM

The main building blocks of the accounts based model are described in this section of the paper. Each subsection deals with a separate component of the accounts and its associated subroutine (named in capitals in the subsection heading). The choices available to the user in constructing his or her accounts based model are outlined with reference to the parameter values that implement the particular options. The outputs produced by each building block of the accounts based model are described. Figure 3.1 shows the sequence of subroutines that make up the heart of the accounts based model. Each subroutine is entered on each iteration of the model, although operations may only be carried out on the first iteration in some subroutines.

3.1 Initial populations (BEGPOP)

Any regional population analysis using ABM requires the input or computation of the numbers of people in each age cohort and sex in each region of interest. There are four sources for the initial populations, as follows.

3.1.1 Initial populations input as data (ipop = 1)

If the ipop parameter is set to 1 (see Tables 4.6 and 4.6.1 for more details), the BEGPOP subroutine reads in the regional populations for an age cohort

$$K_{a}^{e(i) * (*)} \; ; \; i = 1, 2, \ldots, N \qquad 3.1$$

where N is the number of internal regions being studied. This will be done for the A age cohorts $a = 1, \ldots, A$ and, if requested, the X sex groups $x = 1, \ldots, X$ within each cohort

3.1.2 Initial populations are transfered fron the previous period (ipop = 0)

For time intervals other than the first, initial populations for a period are simply set equal to the final populations of the previous time interval for the immediately preceding cohort:

$$K_{a}^{e(i) * (*)} (p_t) = K_{a-1}^{*(*) \, 5(i)} (p_{t-T}) \qquad 3.2$$

where p_t and p_{t-T} are time interval labels for periods beginning at time t and t - T respectively. This option makes possible the projection of the population over many periods, although it is not obligatory in such projections. Fresh or target values of regional populations could be introduced, if required, by reusing the ipop = 1 option.

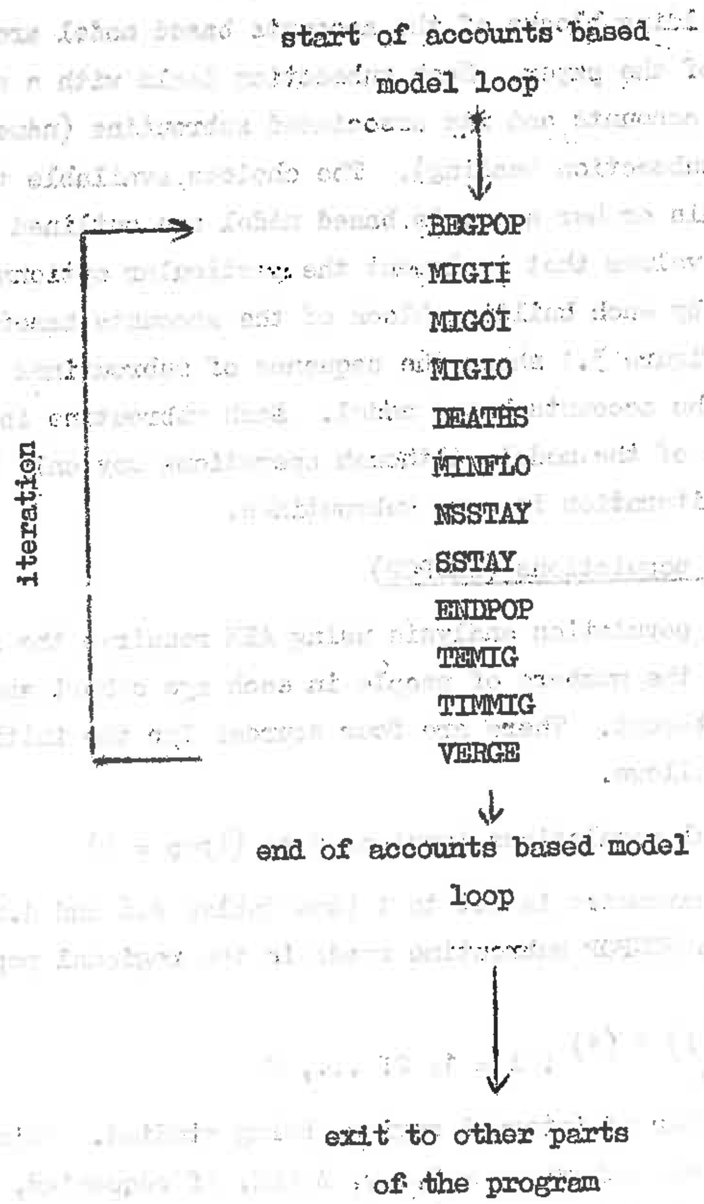


Figure 3.1 The sequence of subroutines in the account based model loop

3.1.3 Births used as initial populations for the infant cohort (ipop = 2)

The "initial populations" in the infant cohort accounts are the births to the regional populations occurring in the various regions:

$$K_0^{b(i)*(*)} \quad , i = 1, 2, \ldots, n \qquad (3.3)$$

These births are worked out in the births model (see section 3.16) and are then used as "initial populations" in the infant age cohort:

$$K_0^{e(i)*(*)} = K^{b(i)*(*)} \qquad (3.4)$$

3.1.4 Initial populations computed from the row accounting equations (ipop = 3)

If initial populations for the regions are unavailable or poorly estimated but final populations are available, then initial populations can be estimated within the accounts based model

$$K_a^{e(i)*(*)} = K_a^{e(i)s(i)} + \sum_{\substack{j \neq i \\ j c I}}^{N} K_a^{e(i)s(j)} + K_a^{e(i)s(R)}$$

$$+ K_a^{e(i)d(i)} + \sum_{\substack{j \neq i \\ j c I}}^{N} K_a^{e(i)s(j)} + K_a^{e(i)d\ (R)} \qquad (3.5)$$

where j c I means region j contained in set I, the internal set of regions, for i = i, ..., N

where $K_a^{e(i)s(i)}$ are surviving stayers in age cohats in region i;

where $\sum_{j \neq i}^{N} K_a^{e(i)s(j)}$ are the artmigrants in age cohort a from region i to all other internal regions; where $K_a^{e(i)s(R)}$ are surviving emigrants in age cohort a from region i to the rest of the world R; and where $K_a^{e(i)d(i)}$, $\sum_{j \neq i}^{N} K_a^{e(i)d(j)}$ and $K_a^{e(i)d(R)}$ are the corresponding terms for non-survivors.

The items in the right hand side of equati on (3.5) are estimated in subsequent steps (or subroutines) of the accounts based model, and the correct choices must be selected for these. In particular, surviving stayers must be computed using the column accounting equation (parameter isstay should be set to 2) and that therefore final populations must be supplied from data (parameter ifpop should be set to 2). Surviving emigrants data should be available (parameter iemig should not be set to 5).

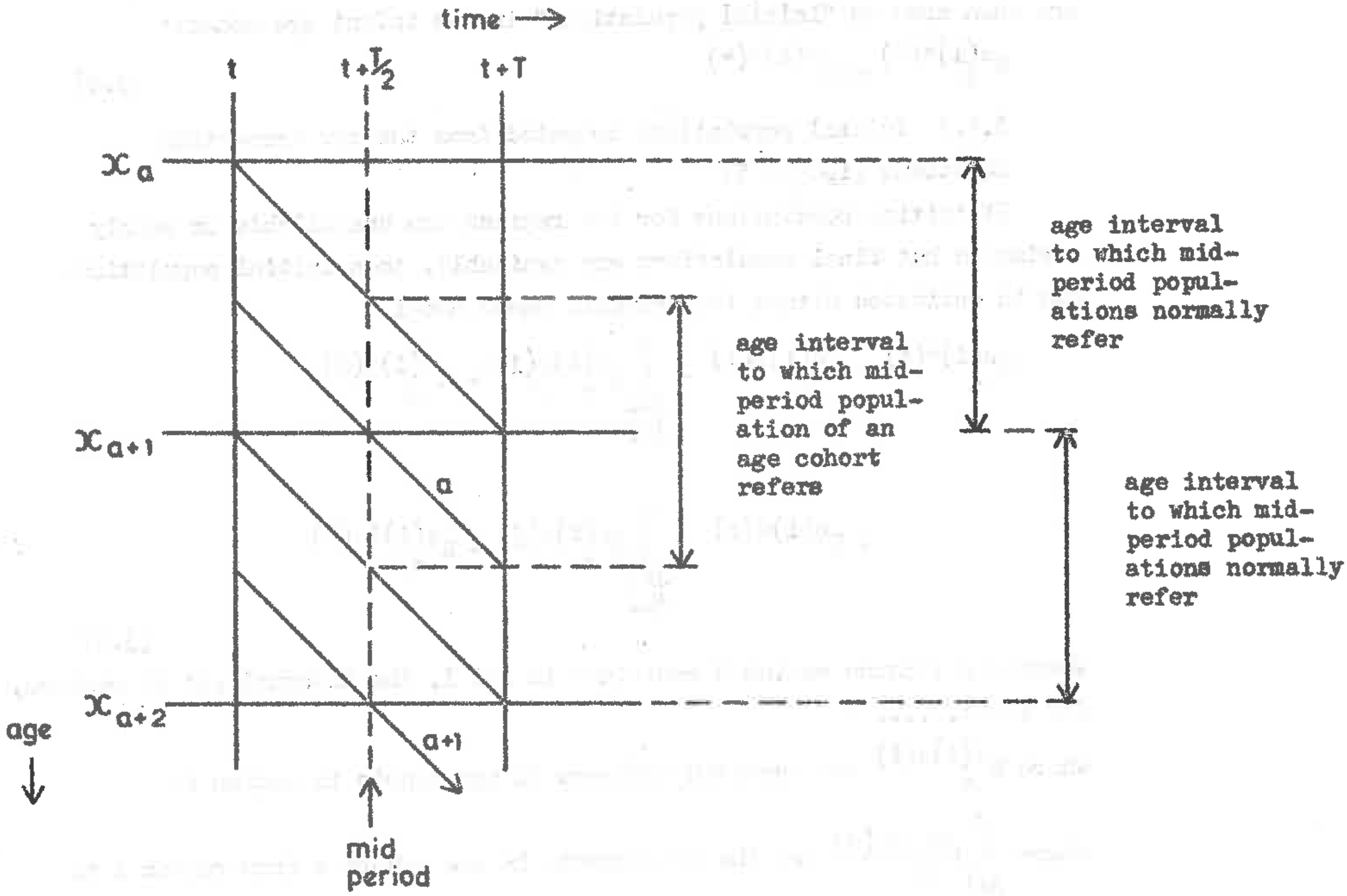


Figure 3.2 A Lexis diagram showing how mid-period and initial and final populations are related for an age cohort

3.1.5 Mid-interval populations only available (ipop = 4)

If only mid-year populations are available these may be input to the program and initial populations are computed using a number of assumptions.

First it is assumed that

$$K_a^{e(i)}(t + {}^T/_2) = \tfrac{1}{2}(K_a^{e(i)}(t + {}^T/_2) + K_{a+1}^{e(i)}(t + {}^T/_2)) \tag{3.6}$$

where $K_a^{e(i)}(t + {}^T/_2)$ is the mid-period population living in region i aged x(a) to x(a) + n where x is the age at which the youngest members of cohort a start the period, and $K_{a+1}^{e(i)}(t + {}^T/_2)$ is the mid-period population living in region i aged x(a + 1) to x(a + 1) + n (that is, in the next age interval). The exact relationships between the age definitions are shown on Figure 3.2's lexis diagram. The estimations involved in equation (3.6) should be carried out prior to inserting mid-period populations into the ABM data file.

Secondly it is assumed that

$$K_a^{e(i)}(t + {}^T/_2) = \tfrac{1}{2}(K_a^{e(i)*(*)}(t) + K_a^{*(x)s(i)}(t + T)\ ) \tag{3.7}$$

that is, that linear change took place in the cohort populations over the the iterval. For the initial populations on the right hand side of equation (3.7) can be substituted all the terms in the equivalent row of accounts matrix (the terms on the right hand side of equation (3.5)), and for the final populations can be substituted the column terms from the account matrix

$$K_a^{*(*)s(i)} = K_a^{e(i)s(i)} + \sum_{\substack{j \neq i \\ j \in I}}^{N} K_a^{e(j)s(i)} + K_a^{e(R)s(i)} \tag{3.8}$$

so that

$$K_a^{e(i)}(t+{}^{T}/_{2}) = \tfrac{1}{2}\ \Big(K_a^{e(i)s(i)} + \sum_{\substack{j\neq i\\ j\in I}} K_a^{e(i)s(j)} + K_a^{e(i)s(R)}$$
$$+ K_a^{e(i)d(i)} + \sum_{\substack{j\neq i\\ j\in I}} K_a^{e(i)d(j)} + K_a^{e(i)d(R)}\Big)$$
$$+ \Big(K_a^{e(i)s(i)} + \sum_{\substack{j\neq i\\ j\in I}} K_a^{e(j)s(i)} + K_a^{e(R)s(i)}\Big) \tag{3.9}$$

This equation is then used to yield an expression for the surviving stayer terms.

$$K_a^{e(i)s(i)} = K_a^{e(i)}(t+{}^{T}/_{2})$$
$$-\tfrac{1}{2}\Big(\sum_{\substack{j\neq i\\ j\in I}} K_a^{e(i)s(j)} + K_a^{e(i)s(R)}$$
$$+ K_a^{e(i)d(i)} + \sum_{j\neq i} K_a^{e(i)d(j)} + K_a^{e(i)d(R)}\Big)$$
$$+ \Big(\sum_{\substack{j\neq i\\ j\in I}} K_a^{e(j)s(i)} + K_a^{e(R)s(i)}\Big) \tag{3.10}$$

This estimate of the surviving stayer terms can be inserted in equation (3.5) to yield estimates of the initial population, the other terms being generated in the subroutines that follow.

Note that if surviving emigrant estimates are not available this option cannot be used. However, if a sequence of three mid-period populations are available initial and final populations can be interpolated :

$$K_a^{e(i)*(*)}(t) = \tfrac{1}{2}(K_{a-1}^{e(i)}(t-{}^{T}/_{2}) + K_a^{e(i)}(t-{}^{T}/_{2}))$$
$$+ \tfrac{1}{2}(K_a^{e(i)}(t+{}^{T}/_{2}) + K_{a+1}^{e(i)}(t+{}^{T}/_{2})) \tag{3.11}$$

$$K_a^{*(*)s(i)}(t+T) = \tfrac{1}{2}\ (K_a^{e(i)}(t+{}^{T}/_{2}) + K_{a+1}^{e(i)}(t+{}^{T}/_{2}))$$
$$+ \tfrac{1}{2}\ (K_{a+1}^{e(i)}(t+{}^{3T}/_{2}) + K_{a+2}^{e(i)}(t+{}^{3T}/_{2})) \tag{3.12}$$

The initial population estimates can then be input to subroutine BEGPOP using parameter ipop set to 1, and final population estimates can be input to subroutine ENDPOP using parameter ifpop set to 2. The iemig parameter must be set to 5 and surviving immigrant values must be available.

3.1.6 Initial population: outputs

Initial populations for the internal regions are reported as vectors, one for each age cohort. Figure 3.3 shows examples of the output for the three region example (East Anglia, South East and Rest of Britain).

```
INITIAL POPULATIONS                      PERIOD =  1   AGE =  1   SEX =  1
----------------------------------------------------------------------------
             EAST ANG   SOUTH EA   REST BRT
              125605.   1409274.   3058065.
TOTAL INITIAL POPULATION =  4592947.
----------------------------------------------------------------------------

INITIAL POPULATIONS                      PERIOD =  1   AGE = 12   SEX =  1
----------------------------------------------------------------------------
             EAST ANG   SOUTH EA   REST BRT
               96590.   1088262.   2202582.
TOTAL INITIAL POPULATION =  3387434.
----------------------------------------------------------------------------
```

Figure 3.3 <u>A printout of initial populations (ipop = 1)</u>

3.2 Surviving internal migrants (MIGII)

3.2.1 Type of migration data

Population accounts for regions require data on the numbers of migrants who move from one region to another and survive there. The persons involved are classified by their initial and final locations; intermediate locations and moves are ignored. In Courgeau's terminology (Courgeau, 1973) accounts involve migrants, not the migrations they make. In Ledent's terminology (Ledent, 1980) accounts are built with data on transitions not data on movements.

Transition or migrant data are conveniently generated by national censuses which include a question about "residence y years ago" or by equivalent retrospective sample surveys. However, in many countries reliance is placed on registration systems which compile statistics on migrations or movements from change of residence forms filled in by citizens who migrate. These migrations data must be converted to migrant data before use in an accounts based model.

This conversion is difficult to do without having available both transition and movement data for the same regional system for the same period of time. For Britain, Ogilvy (1980a) has compared the inter-regional migrant flows recorded in the 1970-71 period by the 1971 Census with the counts of migrations generated from the National Health Service Central Register (NHSCR), and has derived empirical ratios linking two sets of statistics. As both types of migration data will be available for the April 1980-April 1981 year prior to the 1981 Census this should prove to be a very convenient period for calculating such empirical ratios for use with NHSCR figures for the years from 1981 to the next census in 1991.

In the absence of any such comparison simple probability models might be applied to effect the conversion (Rees, 1977; Castro, 1980) but these do not in general work well for interregional migration.

If movements data are used without conversion, the error (the exaggeration of the migration rates) involved will be small (but non-negligible) for one year time intervals (perhaps 10 to 20% on average) but large for five year time intervals (perhaps 30 to 60%). Further work with the British data sets available in 1981 is indicated.

3.2.2 Migrant flows input from the data file (intmig = 1)

The most straightforward way of introducing migrant data into the accounts based model is to supply them as numbers of migrants in each interregional transition:

$$K_a^{e(i)s(j)}, \quad i,j = 1, \ldots, n, \; i \neq j \qquad (3.13)$$

Care must be taken in assigning the migrant numbers to the appropriate cohort as in census tables migrants are classified by age group at the end of the period of measurement. For example, five year migrants recorded in the 10-14 age group in the 1971 census must be assigned to the age cohort that effects the transfer of 5-9 year olds in 1966 to 10-14 year olds in 1971.

This option (intmig = 1) would normally be used in constructing historical accounts (the first row of Figure 2.5) but can also be used to introduce migrant estimates generated externally into projection versions of the accounts based model. Such external models might involve:

(i) generation (of migration) models linked to the projected evolution of the regional economies (the more the unemployment level rises the less migration takes place - see Ogilvy, 1980b);

(ii) gravity models applied to the working ages either in aggregate or in two or three broad groups (Stillwell, 1979, 1980); and the retired ages;

(iii) migration schedule models distributing working age and retired age migration, and generating pre-labour force migrant numbers (following the work of Rogers and Castro, 1978; and Rogers, Raquillet and Castro, 1978).

3.2.3 Migration rates input from the data file (intmig = 3)

Alternatively, migration rates could be input to the accounts based model, and used to generate migrant numbers within the program. This option is particularly useful in projection contexts.

Migration rates are defined as transition rates in the accounts based model, but as movement rates, so that the population at risks assumed are the initial regional populations

$$h_a^{e(i)s(j)} = K_a^{e(i)s(j)} / K_a^{e(i)*(*)} \qquad (3.14)$$

where the letter h is used to denote a transition rate. When transition rates have been read in to the program migrant flows are computed

$$K_a^{e(i)s(j)} = h_a^{e(i)s(j)} \; K_a^{e(i)*(*)} \qquad (3.15)$$

The number of internal migrants remains fixed after the first iteration and the MIGII subroutine is only used for reporting purposes in later iterations.

3.2.4 Migrant flows from the previous period used (intmig = 0)

The very simplest model of the way migration flows evolve over time is to assume they remain constant:

$$K_a^{e(i)s(j)}(p) = K_a^{e(i)s(j)}(p-1) \quad (3.16)$$

where p is the label for the current period and p-1 for the immediately preceding period. This is equivalent to assuming that the net balance of migration remains constant

$$N_a^{e(i)}(p) = N_a^{i}(p-1) \quad (3.17)$$

where

$$N_a^{e(i)}(p-1) = \sum_{\substack{j\neq i \\ j\varepsilon I}} K_a^{e(j)s(i)}(p-1) - \sum_{\substack{j\neq i \\ j\varepsilon I}} K_a^{e(i)s(j)}(p-1) \quad (3.18)$$

This kind of model and assumption is frequently made in current official forecasts for local areas (O.P.C.S., 1978) so that ABM can easily be used to "simulate" official projections. Exact diagnosis of the causes for differences between official and multiregional projection can then be made.

What the program does is to read the required migrant flow values from the direct access file in which are stored the previous period's population accounts values. The user does not, therefore have to input the constant values again.

3.2.5 Migration rates from the previous period used (intmig = 2)

A more likely assumption about migration is that migration rates will remain constant over time

$$h_a^{e(i)s(j)}(p) = h_a^{e(i)s(j)}(p-1) \quad (3.19)$$

Migration rates are then multiplied by the initial cohort populations using equation (3.15) so that migrant flows do fluctuate with the differing cohort sizes over time.

3.2.6 Infant migrant flows generated by a model (intmig = 4)

Some census offices fail to report infant migrant numbers in their migration or place of birth tables so that estimates must be generated for these flows. This is done using a relationship between pre-labour force migration and labour force migration similar to that established by Rogers, Raquillet and Castro (1978):

$$K^{b(i)s(j)} = 0.5\,\Big(\sum_{a=a_1}^{a_2} h_a^{e(i)s(j)f}\; K_a^{b(i)*(*)} \quad (3.20)$$

where $K_a^{b(i)*(*)}$ are the births in region i to parents in age cohort a, $h_a^{e(i)s(j)f}$ are the migration rates from region i to region j for people of the fertile sex f in cohort a and the factor of half is applied because the infants are exposed to the risk of migration for only half the period that their parents are. In practice, we should expect equation (3.20) to exaggerate the infant migrant flow a little as parents are probably slightly less mobile than non-parents in the fertile ages a_1 to a_2.

3.2.7 Surviving internal migrants: outputs

They are reported in the form of a matrix of flows (computed as in equation 3.15 if not input) and a matrix of rates (computed as in equation 3.14 if not input) together with vectors of in-migrant and out-migrant totals (Figure 3.4). Diagonal values, for surviving stayers, are included in the printout for iterations beyond the first.

INTERNAL SURVIVING MIGRANTS PERIOD = 1 AGE = 1 SEX = 1

FLOWS

	EAST ANG 1	SOUTH EA 2	REST BRT 3
EAST ANG	108114.	5739.	6880.
SOUTH EA	11171.	1249325.	58934.
REST BRT	8869.	52833.	2892177.

RATES

	EAST ANG 1	SOUTH EA 2	REST BRT 3
EAST ANG	0.860744	0.045691	0.054775
SOUTH EA	0.007927	0.886503	0.041819
REST BRT	0.002900	0.017277	0.945753

IN-MIGRANT TOTALS

EAST ANG	SOUTH EA	REST BRT
20040.	58572.	65814.

OUT-MIGRANT TOTALS

EAST ANG	SOUTH EA	REST BRT
12619.	70105.	61702.

Figure 3.4 A printout of internal surviving migrants (intmig=1)

3.3 Surviving immigrants (MIGOI)

3.3.1 The nature of immigrant data

Immigrants are, in this paper, defined to be migrants into a region within the system of interest from outside (the external zone). Normally, it will be most convenient to include within the internal set of regions a 'remainder of the country' region (if the internal set does not otherwise exhaust the national territory). Immigrants are then international migrants into a country from abroad. Immigrant numbers are given in some detail in national census migration tables, and most countries also maintain counts of the numbers entering as migrants (planning to stay at least one year). These counts help update immigration statistics between censuses but in certain cases must be treated with reservation either because of their sample nature (International Passenger Survey of the UK) or because of illegal immigration (over the US-Mexico border, for example). Careful comparison of census and count statistics for a period for which statistics from both sources are available is needed (cf Rees, 1979). Counts of immigrations need conversion into estimates of immigrants.

3.3.2 Immigrant flows input from the data file (immig = 1)

Estimates of immigrant flows into the internal regions are read into the program as

$$K_a^{e(R)s(j)} , \quad j = 1, \ldots, n \tag{3.21}$$

when the immig parameter is set to 1.

3.3.3 Immigration rates input from the data file (immig = 3)

Estimates of immigration rates into the internal regions are input to the program and multiplied by populations at risk

$$\hat{K}_a^{e(R)s(i)} = h_a^{e(R)s(i)} \quad \hat{K}_a^{Me(R)s(i)} \tag{3.22}$$

where $\hat{K}_a^{M(R)s(i)}$ is the population of cohort a at risk of immigrating from the external region to region i in the internal system.

3.3.5 Populations at risk of immigration (governed by the iparim parameter)

The populations at risk used to define internal migration rates, namely, the initial populations, are not available in the case of immigrants as the rest of the world populations are assumed not to be known. (If they are known then the 'rest of the world' should be included in the internal set of regions and the immigrant and emigrant flows set to zero).

Instead we use convenient substitutes:

(i) initial populations, $K_a^{e(i)*(*)}$ (iparim = 1)

so that immigration rates are defined as

$$h_a^{e(R)s(i)} = K_a^{e(R)s(i)} \,/\, K_a^{e(i)*(*)} \tag{3.23}$$

(ii) immigration total, $K_a^{e(R)*(*)}$ (iparim = 2)

so that immigration rates are defined as

$$h_a^{e(R)s(i)} = K_a^{e(R)s(i)} / K_a^{e(R)*(*)} \qquad (3.24)$$

in this case, values of the total number of immigrants to the internal system must be separately input; and

(iii) final population, $K_a^{*(*)s(i)}$ (iparim = 3)

so that immigration rates are defined as

$$h_a^{e(R)s(i)} = K_a^{e(R)s(i)} / K_a^{*(*)s(i)} \qquad (3.25)$$

Corresponding with these three rate definitions are three immigrant flow definitions

(i) when iparim = 1

$$K_a^{e(R)s(i)} = h_a^{e(R)s(i)} \quad K_a^{e(i)*(*)} \qquad (3.26)$$

(ii) when iparim = 2

$$K_a^{e(R)s(i)} = h_a^{e(R)s(i)} \quad K_a^{e(R)*(*)} \qquad (3.27)$$

and (iii) when iparim = 3

$$K_a^{e(R)s(i)} = h_a^{e(K)s(i)} \quad K_a^{*(*)s(i)} \qquad (3.28)$$

The second option yields rates which distribute immigrants across the internal regions. These distribution rates are useful in a situation where the total of immigrants into the country may be known but not its distribution within the country. Distribution rates from a previous census-related period could be used with current period totals.

3.3.6 Immigrant flows from the previous period used (immig = 0)

In the absence of better information immigrant flows for the current period can be set equal to those for the previous period

$$K_a^{e(R)s(i)} (p) = K_a^{e(R)s(i)} (p - 1) \qquad (3.29)$$

and this option can be used in cases of projection with constant inputs to the system of interest.

3.3.7 Immigration rates from the previous period used (immig = 2)

Similarly, immigration rates for the current period can be set equal to those of the previous period

$$h_a^{e(R)s(i)} (p) = h_a^{e(R)s(i)} (p-1) \qquad (3.30)$$

Note that in this case the setting of the iparim parameter should be the same in the current period as in the previous.

3.3.8 Infant immigrants estimated (immig = 4)

If no estimates are externally available for the infant cohort, infant immigrants must be estimated within the accounts based model. The estimation model used for infant internal surviving immigrants is not adopted because the births total for the origin region (external region) is not available. Instead infant immigrants are approximated by half the immigrants in the first age cohort.

$$_K b(R)s(i) = 0.5\ K\ {}_1^{}e(R)s(i) \qquad (3.31)$$

3.3.9 Surviving immigrants: outputs

Surviving immigrants are reported as a vector of flows and associated rates as shown in Figure 3.5

SURVIVING IMMIGRANTS		PERIOD = 1	AGE = 1	SEX = 1
FLOWS				
	EAST ANG	SOUTH EA	REST BRT	
	6900.	48268.	58283.	
RATES				
	EAST ANG	SOUTH EA	REST BRT	
	0.051091	0.035592	0.019323	

Figure 3.5 A printout of surviving immigrants

3.4 Surviving emigrants (MIGTO)

3.4.1 The nature of emigrant data

Emigrants are defined as migrants who leave a region within the system of interest to the rest of the world. They are also international migrants if the system of interest includes a "rest of the country" region or exhausts the national territory. No data are available in the census of the country of origin about emigrants, and estimates must be based on count or register or survey data. These population flows are undoubtedly the most difficult to estimate and many countries fail to publish estimates. Nevertheless, many of these countries publish net migration estimates which together with more reliable immigration estimates make possible emigration estimates.

Because of these difficulties an option has been included in the program in which emigrant figures are estimated as residuals given figures on both initial and final populations (as derived in Ledent and Courgeau, 1980). In addition, when some adjustments have to be made to match the total of row constraints on the accounts to the total of column constraints, emigrant totals (surviving and non-surviving) are the most likely to be incorrect and therefore the most obvious to adjust.

3.4.2 Emigrant flows input from the data file (iemig = 1)

Estimates of emigrant flows into the internal regions are read into the program as

$$K_a^{e(i)s(R)} \quad , \quad i = 1 \ , \ \ldots \ , \ n \tag{3.32}$$

when the iemig parameter is set to 1.

Emigration rates from the internal regions are computed using the initial populations as populations at risk

$$h_a^{e(i)s(R)} = K_a^{e(i)s(R)} \ / \ K_a^{e(i)*(*)} \tag{3.33}$$

3.4.3 Emigration rates input from the data file (iemig = 3)

Estimates of emigration rates from the internal regions are read in from the data file as

$$h_a^{e(i)s(R)} \ , \ i = 1 \ , \ \ldots \ , \ n \tag{3.34}$$

and emigrant flows are computed

$$K_a^{e(i)s(R)} = h_a^{e(i)s(R)} \quad K_a^{e(i)*(*)} \tag{3.34}$$

The initial populations are used as populations at risk.

3.4.4 Emigrants flows from the previous period used (iemig = 0)

Emigrant flows for the current period are set equal to those of the previous period

$$K_a^{e(i)s(R)}(p) = K_a^{e(i)s(R)}(p-1) \tag{3.36}$$

Emigration rates are computed using equation (3.33)

3.4.5 Emigration rates from the previous period used (immig = 2)

Emigration rates for the current period are set equal to those of the previous period

$$h_a^{e(i)s(R)}(p) = h_a^{e(i)s(R)}(p-1) \tag{3.37}$$

Emigrant flows are computed using equation (3.34)

3.4.6 Emigrants estimated as residuals from the row accounting equation (iemig = 5)

If emigrant estimates are not available, but if estimates of both initial and final populations are, then surviving emigrants can be estimated using the row accounting equations. This method is most appropriate when the initial and final populations both derive from censuses (Ledent and Courgeau, 1980) but is also appropriate when the populations are only "official" estimates. The residual method can then be regarded as a way of "teasing out" the implicit or explicit emigration figures that went into the final population estimate.

The row accounting equation can be rearranged as follows:

$$K_a^{e(i)s(R)} + K_a^{e(i)d(R)} = K_a^{e(i)*(*)}$$

$$- K_a^{e(i)s(i)} - \sum_{\substack{j\neq i \\ j\, I}} K_a^{e(i)s(j)}$$

$$- K_a^{e(i)d(i)} - \sum_{\substack{j\neq i \\ j\, I}} K_a^{e(i)d(j)}$$

The surviving emigrant and non-surviving terms are grouped on the left hand side of the equation. The right hand side terms are input and computed in other subroutines. Surviving stayers must be computed from the column accounting equations (see subsection 3.9).

Non-surviving emigrants are estimated by the following equation

$$K_a^{e(i)d(R)} = \left(\frac{0.5 \;\; h_a^{*(*)d(i)}}{1 - 0.25 \;\; h_a^{*(*)d(i)}} \right) K_a^{e(i)s(R)} \tag{3.39}$$

(see Rees and Wilson, 1977, Part 2 for the detailed logic underlying this equation where $h_a^{*(*)d(i)}$ is the region i death rate for persons in age cohort a. The right hand side expression can be substituted in the left hand side of equation 3.38 for $K_a^{e(i)d(R)}$ which after rearrangement reads

$$K_a^{e(i)s(R)} = (K_a^{e(i)*(*)} - K_a^{e(i)s(i)} - \sum_{\substack{j \neq i \\ j\, I}} K_a^{e(i)s(j)}$$

$$- K_a^{e(i)d(i)} - \sum_{\substack{j \neq i \\ j\, I}} K_a^{e(i)d(j)})$$

$$\div \left(1 + \frac{0.5\, h_a^{*(*)d(i)}}{1 - 0.25\, h_a^{*(*)d(i)}} \right) \tag{3.40}$$

The right hand side terms are read in or computed in the accounts based model. Although in the first iteration on entry to the MIG 10 subroutine many terms will be still set to zero (all variables are cleared before execution begins - see section 6.4), in subsequent iterations current iteration estimates will be available. The emigrants estimate will be revised until convergence has been achieved.

3.4.7 Surviving emigrants: outputs

Details of the numbers of surviving emigrants from each internal region are reported, if requested, as vectors of flows and associated rates as in Figure 3.6. Flows are computed from equation (3.35) and rates from equation (3.33) if they have not previously been read in.

SURVIVING EMIGRANTS PERIOD = 1 AGE = 1 SEX = 1

	EAST ANG	SOUTH EA	REST BRT
FLOWS	4366.	84055.	89577.
RATES	0.034760	0.059644	0.029292

Figure 3.6 A printout of surviving emigrants

3.5 Deaths (DEATHS, PARI, PARA, and PARM)

The DEATHS subroutine either reads in or computes the number of deceased registered in each internal region, and either computes or reads in the associated death rates under the control of the ipard parameter.

3.5.1 Deaths input from the data file (idth = 1)

If the idth parameter is set to 1, deaths totals for each internal region are read in from the data file as

$$K_a^{*(*)d(i)}, \quad i = 1, \ldots, n \tag{3.41}$$

Death rates are computed by dividing by populations at risk

$$h_a^{*(*)d(i)} = \hat{K}_a^{*(*)d(i)} \;/\; \hat{K}_a^{D*(*)d(i)} \tag{3.42}$$

The population at risk of dying in region, $\hat{K}_a^{D*(*)d(i)}$, is selected by the ipard parameter setting:

(i) if ipard is set equal to 1, initial populations are used

$$\hat{K}_a^{D*(*)d(i)} = \hat{K}_a^{e(i)*(*)} \tag{3.43}$$

(ii) if ipard is set equal to 2, average populations are used

$$\hat{K}_a^{D*(*)d(i)} = \tfrac{1}{2}\,(\hat{K}_a^{e(1)*(*)} + \hat{K}_a^{*(*)s(i)} \tag{3.44}$$

and (iii) if ipard is set equal to 3, multiregional populations at risk are used. These are defined for 'existing' cohorts as

$$\begin{aligned}
\hat{K}_a^{De(*)d(i)} = {} & (1)\,\hat{K}_a^{e(i)s(i)} + (0.5) \sum_{\substack{j\neq i\\ j\epsilon I}} \hat{K}^{e(i)s(j)} \\
& + (0.5)\,K_a^{e(i)s(R)} + (0.5)\,K_a^{e(i)d(i)} + (0.25) \sum_{\substack{j\neq i\\ j\epsilon I}} K_a^{e(i)d(j)} \\
& + (0.25)\,K_a^{e(i)d(R)} + (0.5) \sum_{\substack{j\neq i\\ j\epsilon I}} K_a^{e(j)s(i)} \\
& + (0.5)\,K_a^{e(R)s(i)} + (0.5) \sum_{\substack{j\neq i\\ j\epsilon I}} K_a^{e(j)s(i)} \\
& + (0.5)\,K_a^{e(R)s(i)} + (0.25) \sum_{\substack{j\neq i\\ j\epsilon I}} K_a^{e(j)d(i)} \\
& + (0.25)\,K_a^{e(R)d(i)}
\end{aligned} \tag{3.45}$$

For infant cohorts $\hat{K}_0^{Db(*)d(i)}$ takes the same form with birth or infant terms replacing the existence terms, but multiplied by a further factor of 0.5 reflecting the reduced exposure of infants in a time interval. Equation (3.45) includes all the terms in the region i row and region i columns in the accounts matrix weighted by the proportion of the time interval which persons in each transition spend on average in region i. By implication all the other terms are assigned weights of zero.

Subroutines PARI, PARA and PARM respectively compute the initial population, average population and multiregional populations at risk.

3.5.2 Death rates input from the data file (idth = 3)

If the idth parameter is set to 3, death rates

$$h_a^{*(*)d(i)}, \; i = 1, \ldots, n \qquad (3.46)$$

are read in from the data file and deaths are generated

$$\hat{K}_a^{*(*)d(i)} = h_a^{*(*)d(i)} \; \hat{K}_a^{D*(*)d(i)} \qquad (3.47)$$

where the populations at risk are defined by one of equations (3.43), (3.44) or (3.45) depending on the setting of the ipard parameter. The population at risks selected for use in equation (3.47) should match those used in the definition of the death rates prior to input to ABM.

The population at risk terms in equation (3.42) or those in equation (3.47) will change as fresh estimates of the accounts elements and marginals are made. Deaths rates or death totals for each internal region will be recomputed until convergence in the values of the accounts matrix has occurred.

3.5.3 Deaths from the previous period used (idth = 0)

Deaths for the current period are set equal to those of the previous period

$$K_a^{*(*)d(i)} \, (p) = K_a^{*(*)d(i)} \, (p - 1) \qquad (3.48)$$

Death rates are computed using equation (3.42)

3.5.4 Death rates from the previous period used (idth = 2)

Death rates from the current period are set equal to those for the previous period

$$K_a^{*(*)d(i)} \,)p) = h_a^{*(*)d(i)} \, (p - 1) \qquad (3.49)$$

This option may be useful for constant rate projections or when no further improvements in age specific death rates are forecast.

3.5.5 Deaths computed from column accounting equations (idth = 4)

In certain circumstances, users may not have available deaths data, but have enough survivors data to be able to compute non-surviving stayers via the row accounting equations. Deaths totals can be estimated by adding the columns of the accounts matrix:

$$K_a^{e(*)d(i)} = K_a^{e(i)d(i)} + \sum_{\substack{j\neq i \\ j\in I}} K_a^{e(j)d(i)} + K_a^{e(R)d(i)} \qquad (3.50)$$

and

$$K^{b(*)d(i)} = K_a^{b(i)d(i)} + \sum_{\substack{j\neq i \\ j\in I}} K^{b(j)d(i)} + K^{b(R)d(i)} \qquad (3.51)$$

3.5.6 Deaths: outputs

The deaths total in each internal region are printed out together with the associated populations at risk of dying and the regional death rates as shown in Figure 3.7. The idth parameter was set to 1 (deaths read in from the data file) and the ipard parameter to 3 (multiregional populations at risk computed).

```
POPULATIONS AT RISK                 PERIOD =   1    AGE =   1    SEX =   1
------------------------------------------------------------------------------
                  EAST ANG    SOUTH EA    REST BRT
                   130331.    1382672.    3037181.
------------------------------------------------------------------------------

DEATH TOTALS                        PERIOD =   1    AGE =   1    SEX =   1
------------------------------------------------------------------------------
FLOWS
                  EAST ANG    SOUTH EA    REST BRT
                      530.       5708.      14525.

RATES
                  EAST ANG    SOUTH EA    REST BRT
                  0.004067    0.004128    0.004782
------------------------------------------------------------------------------
```

Figure 3.7 A printout of populations at risk of dying, death totals and death rates

3.6 Non-surviving internal migrants, immigrants and emigrants (MINFLO)

These 'minor flow' items occupy the right hand portion of the accounts matrix for a cohort (Figure 2.2) and must be estimated within the accounts based model because few, if any, demographic statistical systems are sophisticated enough to report them. The models used to generate estimates of the non-survivors sub-matrix are untested as to accuracy: the estimates are, however, fairly well constrained by the column totals (deaths) which are accurately known and by the other terms in the rows of the accounts. Three hypotheses are used in the options for generating the numbers of internal non-surviving migrants; one of the hypotheses has to be used to estimate the number of non-surviving immigrants and another to estimate the number of non-surviving emigrants.

3.6.1 Migrants die at the death rate of their destination region (imin = 1)

In the original design of the accounts based model (Rees and Wilson, 1977), it was hypothesized that migrants die at the rate of the region they migrate to. This leads to the following estimation equation.

$$K_a^{e(i)d(j)} = (0.5\ h_a^{*(*)d(j)} \;/\; (1 - 0.25\ h_a^{*(*)d(j)})\ )\ K_a^{e(i)s(j)} \tag{3.52}$$

If the imin parameter is set to 1, the ABM program uses this equation to compute the numbers of non-surviving internal migrants.

This hypothesis is also used to compute the numbers of non-surviving immigrants

$$K_a^{e(R)d(j)} = (0.5\ h_a^{*(*)d(j)} \;/\; (1 - 0.25\ h_a^{*(*)d(j)})\ )\ K_a^{e(R)s(j)} \tag{3.53}$$

The infant cohort equations are similar but the exposure to risk factors are halved:

$$K^{b(i)d(j)} = (0.25\ h^{*(*)d(j)} / (1 - 0.125\ h^{b(*)d(j)})\ )\ K^{b(i)s(j)} \tag{3.54}$$

and

$$K^{b(R)d(j)} = (0.25\ h^{b(*)d(j)} \;/\; (1 - 0.125\ h^{b(*)d(j)})\ )\ K^{b(R)s(j)} \tag{3.55}$$

3.6.2 Migrants die at the death rate of their origin region (imin = 2)

Alternatively, it could be argued that migrants do not forget their residence histories so easily, and that it would be better to assume that they died at the death rates observed in their origin regions:

$$\hat{K}_a^{e(i)d(j)} = (0.5\ h_a^{*(*)d(i)} \;/\; (1 - 0.25\ h_a^{*(*)d(i)})\ )\ \hat{K}_a^{e(i)s(j)} \tag{3.56}$$

If the imim parameter is set to 2, the program uses this equation to compute non-surviving internal migrant flows.

This hypothesis is also used to compute the numbers of non-surviving emigrants

$$K_a^{e(i)d(R)} = (0.5\, h_a^{*(*)d(i)} \;/\; (1 - 0.25\, h_a^{*(*)d(i)})\;)\; K_a^{e(i)s(R)} \tag{3.57}$$

The infant cohort equations are similar except that the exposure to risk factors are halved (as in (3.52) and (3.53)).

3.6.3 Migrants die at an average of the death rates of their origin and destination regions (inmin = 3)

A compromise hypothesis, reflecting the fact that migrants who die spend some time in both origin and destination regions, would be to assume that migrants die at some average of their origin and destination region death rates. If the arithmetic average is selected then non-surviving internal migrants are estimated by

$$K_a^{e(i)d(j)} = (0.5\, (h_a^{*(*)d(i)} + h_a^{*(*)d(j)}) \;/\; (1 - 0.25\, (h_a^{*(*)d(i)} + h_a^{*(*)d(j)})))\; \times\; K_a^{e(i)s(j)} \tag{3.58}$$

This equation is used if the imin parameter is set to 3.

3.6.4 Non-surviving migrants: outputs

Figure 3.8 shows printouts of non-surviving internal migrants, in- and out-migrant totals for non-survivors and the non-surviving emigrants and immigrants for East Anglia, the South East and the Rest of Britain for the 0-4 to 5-9 cohort in 1966-71. The imin parameter was set to 1: the destination region death rates were used.

The rates are computed using the initial populations as populations at risk in the case of non-surviving internal migrants.

$$h_a^{e(i)d(j)} = K_a^{e(i)d(j)} / K_a^{e(i)*(*)} \tag{3.59}$$

$$h_a^{e(i)d(R)} = K_a^{e(i)d(R)} / K_a^{e(i)*(*)} \tag{3.60}$$

with equivalent definitions for the infant cohort replacing superscript 'e' by 'b', and in the case of non-surviving immigrants the rates are defined using populations at risk specified by the iparim parameter (see sub-section 3.3)

$$h_a^{e(R)d(i)} = K_a^{e(R)d(i)} / K_a^{M\; e(R)d(i)} \tag{3.60}$$

```
NON-SURV. INTERNAL MIGRANTS     PERIOD =  1    AGE =  1    SEX =  1
------------------------------------------------------------------------
FLOWS
                 EAST ANG   SOUTH EA   REST BRT
                        1          2          3
EAST ANG               0.        12.        16.
SOUTH EA              23.         0.       141.
REST BRT              18.       109.         0.

RATES
                 EAST ANG   SOUTH EA   REST BRT
                        1          2          3
EAST ANG         0.0        0.000094   0.000131
SOUTH EA         0.000016   0.0        0.000100
REST BRT         0.000006   0.000036   0.0
------------------------------------------------------------------------

IN-MIGRANT INTERNAL TOTALS FOR NON-SURVIVORS
                                PERIOD =  1    AGE =  1    SEX =  1
------------------------------------------------------------------------
                 EAST ANG   SOUTH EA   REST BRT
                      41.       121.       158.

OUT-MIGRANT INTERNAL TOTALS FOR NON-SURVIVORS
                                PERIOD =  1    AGE =  1    SEX =  1
------------------------------------------------------------------------
                 EAST ANG   SOUTH EA   REST BRT
                      28.       164.       127.

NON-SURVIVING EMIGRANTS         PERIOD =  1    AGE =  1    SEX =  1
------------------------------------------------------------------------
FLOWS
                 EAST ANG   SOUTH EA   REST BRT
                       9.       174.       214.
RATES
                 EAST ANG   SOUTH EA   REST BRT
                 0.000071   0.000123   0.000070

NON-SURVIVING IMMIGRANTS        PERIOD =  1    AGE =  1    SEX =  1
------------------------------------------------------------------------
FLOWS
                 EAST ANG   SOUTH EA   REST BRT
                      14.       106.       140.
RATES
                 EAST ANG   SOUTH EA   REST BRT
                 0.000112   0.000071   0.000046
------------------------------------------------------------------------
```

Figure 3.8 A printout of non-surviving migrants

3.7 Non-surviving stayers (NSSTAY)

3.7.1 Non-surviving stayers computed from the column accounting equations (instay = 1)

The diagonal terms in the non-survivors part of the accounts matrix, the non-survivors who die in the same region that they started the period in, are estimated in the original accounts based model (Rees and Wilson, 1977) from the column accounting equation by subtracting from the known or computed deaths total the total of non-surviving in-migrants.

$$K_a^{e(i)d(i)} = K_a^{e(*)d(i)} - \sum_{\substack{j\neq i\\ j\in I}} K_a^{e(j)d(i)} - K_a^{e(R)d(i)} \qquad (3.61)$$

for existing cohorts and

$$K^{b(i)d(i)} = K^{b(*)d(i)} - \sum_{\substack{j\neq i\\ j\in I}} K^{b(j)d(i)} - K^{b(R)d(i)} \qquad (3.62)$$

for the infant cohort. These equations are used when the instay parameter is set equal to 1, and will be the option normally used.

3.7.2 Non-surviving stayers computed from the row accounting equation (instay = 2)

Occasionally, it may be more convenient to compute the number of non-surviving stayers from the row accounting equations thus

$$K_a^{e(i)d(i)} = K_a^{e(i)*(*)} - K_a^{e(i)s(i)} - \sum_{\substack{j\neq i\\ j\in I}} K_a^{e(i)s(j)} - K_a^{e(i)s(R)} - \sum_{\substack{j\neq i\\ j\in I}} K^{e(i)d(j)} - K_a^{e(i)d(R)} \qquad (3.63)$$

with there being an equivalent equation for the infant cohort. This option will be associated with the computation of the regional deaths total from the column accounting equation (idth = 4).

3.7.3 Non-surviving stayers computed from a model equation (instay = 3)

Just as non-surviving migrants were computed using model equations so can non-surviving stayers be so computed. The estimating equation is derived as follows:

$$\hat{K}_a^{e(i)d(i)} = \hat{n}_a^{*(*)d(i)} \quad \hat{K}_a^{De(i)d(i)} \qquad (3.64)$$

where $\hat{K}_a^{De(i)d(i)}$ is the population at risk of dying in the same region i as it began the period and is defined as

$$K_a^{De(i)d(i)} = (1)\ K_a^{e(i)s(i)} + (0.5) \sum_{\substack{j \neq i \\ j \in I}} K_a^{e(i)s(j)} + (0.5)\ K_a^{e(i)s(R)} + (0.5)\ K_a^{e(i)d(i)} + (0.25) \sum_{\substack{j \neq i \\ j \in I}} K^{e(i)d(j)} + (0.25)\ K_a^{e(i)d(R)} \quad (3.65)$$

If the RH side of (3.65) is substituted into (3.64) and the resulting equation is rearranged, we obtain

$$K_a^{e(i)d(i)} = (h_a^{*(*)d(i)} / (1 - (0.5)\ h_a^{*(*)d(i)})) \times (K_a^{e(i)s(i)} + (0.5) \sum_{\substack{j \neq i \\ j \in I}} K^{e(i)s(j)} + (0.5)\ K_a^{e(i)s(R)} + (0.25) \sum_{\substack{j \neq i \\ j \in I}} K_a^{e(i)d(j)} + (0.25)\ K_a^{e(i)d(R)}) \quad (3.66)$$

The infant cohort equation is similar except that the right side is multiplied by one half.

Total deaths figures may also be available when this option is used. They can be introduced at a later stage by adjusting the accounts matrix to known constraints (see subsection 3.14).

3.7.4 Non-surviving stayers: outputs

These are reported in the familiar form of tables of flows and associated rates, computed using the initial populations as populations at risk (Figure 3.9). Non-surviving stayers are here computed using the column accounting equation (instay = 1).

NON-SURVIVORS INC. STAYERS (1) PERIOD = 1 AGE = 1 SEX = 1

FLOWS

	EAST ANG 1	SOUTH EA 2	REST BRT 3
EAST ANG	475.	12.	16.
SOUTH EA	23.	5487.	141.
REST BRT	18.	199.	14228.

RATES

	EAST ANG 1	SOUTH EA 2	REST BRT 3
EAST ANG	0.003783	0.000094	0.000131
SOUTH EA	0.000016	0.003894	0.000100
REST BRT	0.000006	0.000036	0.004653

Figure 3.9 <u>A printout of non-survivors including stayers for internal regions (instay = 1)</u>

SURVIVORS INC. STAYERS PERIOD = 1 AGE = 1 SEX = 1

	EAST ANG 1	SOUTH EA 2	REST BRT 3
EAST ANG	108108.	5739.	6880.
SOUTH EA	11171.	1249289.	58934.
REST BRT	8869.	52833.	2892219.

RATES

	EAST ANG 1	SOUTH EA 2	REST BRT 3
EAST ANG	0.860695	0.045691	0.054775
SOUTH EA	0.007927	0.886477	0.041819
REST BRT	0.002900	0.017277	0.945767

Figure 3.10 <u>A printout of internal region survivors including stayers (instay = 1)</u>

3.8 Surviving stayers (SSTAY)

3.8.1 Surviving stayers computed using the row accounting equations (isstay = 1)

At this stage of the accounts based model all terms in the rows of the accounts table have been estimated except for the surviving stayer terms. These are estimated as residuals from the row accounting equations.

$$K_a^{e(i)s(i)} = K_a^{e(i)*(*)} - \sum_{\substack{j\neq i \\ j\epsilon I}} K_a^{e(i)s(j)} - K_a^{e(i)s(R)}$$
$$- K_a^{e(i)d(i)} - \sum_{\substack{j\neq i \\ j\epsilon I}} K_a^{e(i)d(j)} - K_a^{e(i)d(R)} \qquad (3.67)$$

with a similar equation for the infant cohort. These equations are used when the isstay parameter is set to 1.

3.8.2 Surviving stayers computed using the column accounting equations (isstay = 2)

In the earlier discussion about the nature of the migrant data demanded by the accounts based model, the principal source of the migrant statistics needed was identified as the retrospective migration tabulations of the periodic national censuses. If the response to the question 'where were you living y years ago?' was 'the same place or region as now', this would identify 'surviving stayers'. Normally, these 'diagonal' terms are not published in the census tabulations (more usually the diagonal terms are the numbers of migrants who have migrated within a region), but they are easily computed if the final populations recorded at the census are input as well (see section 3.9). The column accounting equations are used

$$K_a^{e(i)s(i)} = K^{*(*)s(i)} - \sum_{\substack{j\neq i \\ j\epsilon I}} K^{b(j)s(i)} - K_a^{e(R)s(i)} \qquad (3.68)$$

and for infants

$$K^{b(i)s(i)} = K^{(*)s(i)} - \sum_{\substack{j\neq i \\ j\epsilon I}} K^{b(j)s(i)} - K^{b(R)s(i)} \qquad (3.69)$$

If the isstay parameter is set to 2, these equations are used.

3.8.3 Surviving stayers: outputs

Surviving stayers are printed out in tabular form along with surviving internal migrants input or computed earlier in the sequence of accounts based model subroutines, together with the association transition rates computed using initial populations as populations at risk (Figure 3.10).

```
FINAL POPULATIONS              PERIOD =  1   AGE =  1   SEX =  1
------------------------------------------------------------------
             EAST ANG   SOUTH EA   REST BRT
             135048.    1356129.   3016316.
------------------------------------------------------------------
```

Figure 3.11 A printout of final populations and population change rates (ifpop = 1)

```
POPULATION CHANGE RATES        PERIOD =  1   AGE =  1   SEX =  1
------------------------------------------------------------------
             EAST ANG   SOUTH EA   REST BRT
             1.075177   0.962289   0.986347
------------------------------------------------------------------
```

```
EMIGRANT TOTALS                PERIOD =  1   AGE =  1   SEX =  1
------------------------------------------------------------------
TOTAL SURVIVING EMIGRANTS            =   177998.
TOTAL NON-SURVIVING EMIGRANTS        =      597.
TOTAL SURVIVING EMIGRATION RATE      =  0.038755
TOTAL NON-SURVIVING EMIGRATION RATE  =  0.000086
------------------------------------------------------------------
```

Figure 3.12 A printout of emigrant totals and associated rates (iemig = 1)

```
IMMIGRANT TOTAL                PERIOD =  1   AGE =  1   SEX =  1
------------------------------------------------------------------
FLOW =     113704.                 RATE =  0.024756
```

Figure 3.13 A printout of the immigrant total flow and associated rate (immig = 1)

3.9 Final populations (ENDPOP)

3.9.1 Final populations computed from the column accounting equations (ifpop = 1)

The crucial estimate for most users produced by the accounts based model is of final populations. These are estimated using the appropriate column accounting equations

$$K_a^{e(*)s(i)} = K_a^{e(i)s(i)} + \sum_{\substack{j \neq i \\ j \in I}} K^{e(j)s(i)} + K^{e(R)s(i)} \tag{3.70}$$

and for the infant cohort

$$K^{b(*)s(i)} = K^{b(i)s(i)} + \sum_{\substack{j \neq i \\ j \in I}} K^{b(j)s(i)} + K^{b(R)s(i)} \tag{3.71}$$

These equations are used when the ifpop parameter is set to 1 which will normally be the case when unconstrained accounts are computed (Figure 2.5), particularly in 'estimate' or 'projection' mode.

3.9.2 Final populations input from the data file (ifpop = 2)

Frequently, however, final populations are known when historical accounts are prepared: they derive from the same census that yields the initial internal migrant and immigrant data. So, a second option (the ifpop parameter set to 2) allows the variables

$$K_a^{e(*)s(i)}, \; i = 1, \ldots, n \tag{3.72}$$

$$K^{b(*)s(i)}, \; i = 1, \ldots, n \tag{3.73}$$

to be read in from the data file. This option should logically be used with isstay set equal to 2, and if initial populations are also known and read in (ipop = 1) the accounts should be adjusted to constraints to ensure consistency.

3.9.3 Final populations: outputs

Final populations of the internal regions are output in vector form and rates of population change are computed by dividing final populations in a cohort by initial. Figure 3.11 shows the printout produced by the program. The first cohort, born in the 1961-66 period, increases in size in East Anglia while in the South East and the Rest of Britain the first cohort decreases in the 1966-71 period - at the end of the period members of this cohort are aged 5-9.

3.10 Emigrant totals (TEMIG)

Emigrant totals are computed from the estimates made in the accounts based model of surviving and non-surviving emigrants. Total surviving emigrants are given by

$$K_a^{e(*)s(R)} = \sum_{i \in I} K_a^{e(i)s(R)} \tag{3.74}$$

and

$$K^{b(*)s(R)} = \sum_{i \in I} K^{b(i)s(R)} \tag{3.75}$$

recalling that $K_a^{e(R)s(R)}$ and $K^{b(R)s(R)}$ have been set to zero.

Total non-surviving emigrants are similar sums

$$K^{e(*)d(R)} = \sum_{i \in I} K_a^{e(i)d(R)} \tag{3.76}$$

and

$$K^{b(*)d(R)} = \sum_{i \in I} K^{b(i)d(R)} \tag{3.77}$$

Figure 3.12 shows the printout of emigrant totals for an age cohort. The total emigration rates are computed by dividing the surviving and non-surviving emigrant totals by the total initial population of the internal regions

$$h_a^{e(*)s(R)} = K_a^{e(*)s(R)} \;/\; \sum_{i \in I} K_a^{e(i)*(*)} \tag{3.78}$$

$$h_a^{e(*)d(R)} = K_a^{e(*)d(R)} \;/\; \sum_{i \in I} K_a^{e(i)*(*)} \tag{3.79}$$

with equivalent definitions for the infant cohort.

3.11 Immigrant total (TIMMIG)

The total number of immigrants to the set of regions constituting the system of interest is computed as the sum of the terms in the external region row in each cohort's accounts

$$K_{a}^{e(R)*(*)} = \sum_{j \in I} K_{a}^{e(R)s(j)} + \sum_{j \in I} K_{a}^{e(R)d(j)} \tag{3.80}$$

with an equivalent equation for the infant cohort.

However, if, in connection with the surviving immigrant component, the second option was selected to use distributive rates and an immigration total read in from the data file, this independent total remains the value for $K^{e(R)*(*)}$. In this case adjustment of the accounts matrix via the constraints and balancing factor routine will be appropriate.

Figure 3.13 shows the printout of the immigrant total for an age cohort. The total immigration rate is computed by dividing the immigration total by the total population of the system of interest

$$h_{a}^{(R)*(*)} = K^{e(R)*(*)} \;/\; \sum_{i \in I} K^{e(i)*(*)} \tag{3.81}$$

3.12 Assembly of the accounts (ACCASS,ACCNTS)

Once all the components or building blocks of the accounts have been input or computed they are assembled into an accounts matrix in the ACCASS and ACCNTS subroutines and printed out in table format as shown in Figure 3.14. The rates associated with the accounts shown in the second table are those which have been described for each component. The internal region rows all contain transition rates in which the initial populations are the divisors: hence the total rates are all 1.0. The external region row contains immigration rates, divisors of which are defined by the iparim parameter choice. The total immigration rate has the total internal region population as its divisor. The rates in the final population position are rates of change of population; the rates in the deaths columns for the internal region are death rates defined by the ipard parameter choice. The emigrant total rates have the total internal region population as divisor.

These two tables of accounts and associated rates contain for a cohort all the information used and generated by the accounts based model. Previous printouts are therefore in part redundant and users may suppress the printout of all components other than final populations in all periods beyond the first when initial populations are also printed.

3.13 Components of growth (COG)

An alternative and more familiar way to report on population change is to identify the "components of growth" on a single region basis. This is done in the next table printed out by the ABM program. The definitions of most of the components are clear from the column headings and table notes. Users should note that the components apply to age cohorts; that for all cohorts bar the infant cohort the births column is blank; that the natural increase column is also blank; and that all of the migration components include non-surviving migrants as well as surviving migrants.

THE ACCOUNTS MATRIX			PERIOD = 1	AGE = 1	SEX = 1	ITERATION = 3			
	EAST ANG 1	SOUTH EA 2	REST BRT 3	REST WLD 4	EAST ANG 5	SOUTH EA 6	REST BRT 7	REST WLD 8	TOTALS 9
EAST ANG	108108.	5739.	6880.	4366.	475.	12.	16.	9.	125615.
SOUTH EA	11171.	1249289.	58934.	84055.	23.	5487.	141.	174.	1409274.
REST BRT	8869.	52833.	2892219.	89577.	18.	109.	14228.	214.	3058068.
REST WLD	6900.	48268.	58283.	0.	14.	130.	140.	0.	113704.
TOTALS	135048.	1356129.	3016316.	177998.	530.	5738.	14525.	397.	4706551.

THE RATES MATRIX ASSOCIATED WITH THE ACCOUNTS

	EAST ANG 1	SOUTH EA 2	REST BRT 3	REST WLD 4	EAST ANG 5	SOUTH EA 6	REST BRT 7	REST WLD 8	TOTALS 9
EAST ANG	0.860695	0.045691	0.054775	0.034760	0.003783	0.000094	0.000131	0.000071	1.000000
SOUTH EA	0.007927	0.886477	0.041819	0.059644	0.000016	0.003894	0.000100	0.000123	1.000000
REST BRT	0.002900	0.017277	0.945767	0.029292	0.000006	0.000036	0.004653	0.000070	1.000000
REST WLD	0.054934	0.134250	0.019059	0.0	0.000112	0.000071	0.000046	0.0	0.024756
TOTALS	1.075177	0.962289	0.986347	0.038755	0.004067	0.004128	0.004782	0.000086	1.000000

Figure 3.14 The accounts matrix and associated rates matrix for the first cohort in East Anglia, the South East and Rest of Britain, 1966-71

COMPONENTS OF GROWTH PERIOD = 1 AGE = 1 SEX = 1 ITERATION = 3

REGION	FINAL POPULATION (1)	INITIAL POPULATION (2)	POPULATION CHANGE (3)	BIRTHS (4)	DEATHS (5)	NATURAL INCREASE (6)
EAST ANG	135048.	125605.	9443.		530.	
SOUTH EA	1356129.	1409274.	-53145.		5708.	
REST BRT	3016316.	3058068.	-41752.		14525.	

REGION	INTERNAL IN-MIGRATION (7)	INTERNAL OUT-MIGRATION (8)	INTERNAL NET MIGRATION (9)	EXTERNAL IN-MIGRATION (10)	EXTERNAL OUT-MIGRATION (11)	EXTERNAL NET MIGRATION (12)	TOTAL IN-MIGRATION (13)	TOTAL OUT-MIGRATION (14)	TOTAL NET MIGRATION (15)
EAST ANG	20081.	12647.	7433.	6914.	4375.	2539.	26995.	17022.	9973.
SOUTH EA	58693.	70269.	-11576.	48368.	84229.	-35861.	107061.	154498.	-47437.
REST BRT	65972.	61829.	4142.	58423.	89791.	-31369.	124394.	151621.	-27227.

NOTES ON THE COMPONENTS OF GROWTH

COLUMN 3 = 1 - 2 COLUMN 6 = 4 - 5 COLUMN 9 = 7 - 8 COLUMN 12 = 10 - 11

COLUMN 15 = 13 - 14 COLUMN 3 = 6 + 15 COLUMN 13 = 7 + 10 COLUMN 14 = 8 + 11

Figure 3.15 The components of growth for the first cohort in East Anglia, South East and Rest of Britain, 1966-71

3.14 Testing for convergence (VERGE)

As was indicated earlier (see Figure 3.1) the accounts based model has a loop structure: after total immigrants have been computed, the program returns to the initial populations subroutine, BEGPOP, and recomputes terms that need recomputing. Frequently, only on the second iteration of the accounts based model will a reasonable set of accounts be estimated.

How many times should the model calculations be repeated? The model is iterated until the accounts estimates cease to change from one iteration to the next by more than a negligible amount. The most convenient way to test whether such a convergence has taken place is to compare values of the multiregional populations at risk at successive iterations. The multiregional populations at risk are a weighted function of all elements in the accounts matrices. Does the absolute value of the expression

$$(\hat{K}_a^{De(*)d(i)}(k) - \hat{K}_a^{De(*)d(i)}(k-1))/(\hat{K}_a^{De(*)d(i)}(k-1)) \qquad (3.82)$$

exceed a fixed small value or not? When this question is answered negatively for all regions then the accounts will have converged. The fixed small value must be selected by the user, but a number in the range .0001 or less is recommended to achieve convergence in the last digit of a population figure of the magnitude in the three region example.

At each iteration the multiregional populations at risk are written out and a message is printed informing the user of whether convergence has been achieved or not (see Figure 3.16). In the example shown in Figure 3.16 convergence has been achieved at the third iteration at a tolerance level of .0001 or 1 in 10,000. All values of the multiregional population at risk at the third iteration differ by less than 1 in 10,000 from the values at the previous iteration. In fact, this tolerance level does not guarantee that the populations at risk have converged to single person accuracy and should be revised to .00001 or .000001.

```
------------------------------------------------------------------------------------------
POPULATIONS AT RISK            PERIOD =  1    AGE =  1    SEX =  1    ITERATION =  3
------------------------------------------------------------------------------------------
          EAST ANG   SOUTH EA   REST BRT
           130331.   1382672.   3037181.
------------------------------------------------------------------------------------------
IFLAG =  0    CONVERGENCE ACHIEVED
```

Figure 3.16 A printout of multiregional populations at risk and the convergence achieved/not achieved message

3.15 Constrained accounts (CONSTR)

3.15.1 General Discussion

In constructing the accounts based model the user may have experienced an 'embarras de richesses': more data than were necessary to make an estimate of the accounts or alternative estimates of some of the marginal totals generated by the accounts based model may be available. In this situation the user assembles estimates for the row and column marginals of the accounts and the program (subroutine CONSTR) adjusts the accounts matrix to satisfy these constraints. If no adjustment is required then the icon parameter should be set to zero; otherwise a value of 1, 2, or 3 should be adopted.

3.15.2 A balancing factor model

A simple two balancing factor model is used. Let us simplify the accounts notation for each cohort to K_a^{mn} with m the label for rows and n the column label, and call the row constraints, R_a^m, and the column constraints, C_a^n. The first n row constraints will be the initial populations of the regions; the (n + 1)th row constraint will be total number of immigrants. The first n column constraints will refer to the final populations of the regions; the (n + 1)th column constraint is the total of surviving emigrants; the next n column constraints will be the total deaths in the internal regions; the final column constraint will be the number of non-surviving emigrants.

The adjusted accounts are derived as follows:

$$\hat{K}_a^{mn} = A_a^m \; B_a^n \; \hat{K}_a^{mn} \qquad (3.83)$$

subject to

$$\sum_n \hat{K}_a^{mn} = R_a^m \qquad (3.84)$$

and

$$\sum_m \hat{K}_a^{mn} = C_a^n \qquad (3.85)$$

substituting the right hand side of equation (3.83) in the left hand sides of equations (3.84) and (3.85), and rearranging, we obtain

$$A_a^m = R_a^m / \sum_n B_a^n \; K_a^{mn} \qquad (3.86)$$

$$B_a^u = C_a^n / \sum_m A_a^m \; K_a^{mn} \qquad (3.87)$$

These equations must be solved iteratively.

The row and column constraints are checked to see whether they add up to the same total. For the accounts to be successfully adjusted it is essential that

$$\sum_m R_a^m = \sum_n C_a^n \qquad (3.88)$$

Because row and column constraints are derived from many different statistical sources there is likely to be a difference between the two totals. This difference must be distributed among the row and column constraints to ensure that equation (3.88) is satisfied. Which terms will be loaded with the difference will depend very much on the population system being studied and the demographic sources being used, so that these adjustments are left outside the program.

Then the iterative procedure is carried out as follows. The B^n_a terms are set equal to 1 and the A^m_a's are computed using equation (3.86). The B^n_a balancing factors are then computed using equation (3.87). After this is repeated sufficient times the values of the balancing factors from one iteration differs from the next by less than a very small amount. Convergence of the process may be tested for in one of three ways.

3.15.3 Convergence criteria

(i) Relative differences of the balancing factors (icon = 1)

If for all m, m = 1, ..., n + 1 and all n, n = 1, ..., 2n + 2, the following conditions hold

$$\text{abs}\ ((A^m_a(k) - A^m_a(k-1)) / A^m_a(k-1)) < \text{tol} \qquad (3.89)$$

and

$$\text{abs}\ ((B^n_a(k) - B^n_a(k-1)) / B^n_a(k-1)) < \text{tol} \qquad (3.90)$$

where abs ( ) refers to the absolute value of the function in parentheses, where tol is a small decimal number such as .000001 or .0000001, and k-1 and k are successive iteration numbers, then the process is held to have converged. Equation (3.83) is then used to adjust the accounts matrix with the converged balancing factors. This criterion for convergence is adopted when the icon parameter is set equal to 1.

(ii) Relative differences of the accounts elements (icon = 2)

If the icon parameter is set to 2, the program examines the relative differences of accounts elements at successive iterations, and if

$$\text{abs}\ ((\hat{K}^{mn}_a(k) - \hat{K}^{mn}_a(k-1)) / \hat{K}^{mn}_a(k)) < \text{tol} \qquad (3.91)$$

for all m and n, then the accounts are said to have converged. Tol will be a small number in the order of 10^{-6} to 10^{-9}.

(iii) Absolute differences of the accounts elements (icon = 3)

A more stringent criterion for convergence of the accounts adjustment procedure is to test whether the differences between accounts elements at successive iterations. If for all m and all n

$$\text{abs}\ (\hat{K}^{mn}_a(k) - \hat{K}^{mn}_a(k-1)) < \text{tol} \qquad (3.92)$$

then the adjusted accounts are accepted as having converged. This criterion is adopted by the program if icon is set equal to 3.

For each criterion the user needs to specify the tolerance level or 'tol' for the differences beteen successive iterations which he or she is prepared to accept. For the first two criteria tol should be set to a small number in the range 10^{-6} to 10^{-9}. For the third criterion, a number such as .49 will ensure that accounts elements differ from each other at successive iterations by less than one person.

3.15.4 Adopting accounts elements as constraints (iconk = 1)

The process described above adjusts the accounts matrix for an age cohort to a pre-determined set of marginals. However, users may wish to hold constant in the adjustment procedure one or other of the elements of the accounts matrix. To do this users must set the iconk parameter to 1 and include a matrix of 0 and 1's in the data file as specified below; otherwise the iconk parameter should be assigned a value of zero.

This may be accomplished by the following modification of the balancing factor model. A matrix, F, of 0's and 1's of dimensions (n + 1) x (2n + 2) corresponding to the accounts matrix for a cohort, $\underline{K}_a$, is input by the user: a zero in the matrix F indicates that the element in the corrrsponding position in the accounts matrix is to remain fixed; a one indicates that the accounts element in the corresponding position is freed to be adjusted. Each element in the K_a matrix is multiplied by the corresponding element in the $\underline{F}$ matrix to form a new accounts matrix, $\underline{K}_a(F)$ for adjustment:

$$\underline{K}_a(F) = \underline{F} \otimes \underline{K}_a \tag{3.93}$$

where $\otimes$ indicates the matrix operation described above in which

$$K_a^{mn}(F) = F^{mn} K_a^{mn} \tag{3.94}$$

Let us denote the row and column sums of $\underline{J}_a$ and $\underline{K}_a$ as

$$K_a^{m*}(F) = \Sigma_n K_a^{mn}(F)$$

$$K_a^{*n}(F) = \Sigma_m K_a^{mn}(F)$$

$$K_a^{m*} = \Sigma_n K_a^{mn}$$

$$K_a^{*n} = \Sigma_m K_a^{mn} \tag{3.95}$$

Then the differences

$$K_a^{m*} - K_a^{m*}(F)$$

$$K_a^{*n} - K_a^{*n}(F)$$

represent the row and column sums of the fixed elements. These differences are subtracted from the row and column constraints to yield revised row and column constraints

$$R_a^m(F) = R_a^m - (K_a^{m*} - K_a^{m*}(F)) \tag{3.96}$$

$$C_a^n(F) = C_a^n - (K_a^{*n} - K_a^{*n}(F)) \tag{3.97}$$

and a new balancing factor model is computed in which the accounts matrix with fixed elements set to zero is substituted for the original accounts matrix and the revised constraints are employed

$$\hat{K}_a^{mn}(F) = A_a^m(F)\, B_a^n(F)\, K_a^{mn}(F) \tag{3.98}$$

subject to

$$\sum_n \hat{K}_a^{mn}(F) = R_a^m(F) \tag{3.99}$$

and

$$\sum_m \hat{K}_a^{mn}(F) = C_a^n(F) \tag{3.100}$$

The fixed elements do not affect the accounts balancing procedure, having been set to zero.

To obtain the final revised accounts the $\underline{F}$ matrix is 'reversed' (elements that were 0 are set to 1 and elements that were 1 are set to zero) to form a matrix $\underline{G}$ and the elements of this are multiplied by the corresponding elements of $\underline{K}_a$

$$\underline{K}_a(G) = \underline{G} \otimes \underline{K}_a \tag{3.101}$$

The revised, adjusted accounts matrix with fixed elements is then

$$\hat{\underline{K}}_a = \hat{\underline{K}}_a(F) + \underline{K}_a(G) \tag{3.102}$$

Convergence of the accounts adjustment procedure with fixed elements depends on there being at least one element in each row and each column of the accounts which is not fixed, and convergence in a reasonable number of iterations probably depends on the number of fixed elements being restricted in number.

3.15.5 Constraints procedure: outputs

If requested the program will printout details of all the iterations required to balance the accounts matrix to the constraints. Figure 3.17 shows the details for the last iteration.

3.16 The births model

3.16.1 General features

Once the ABM program has computed accounts for all age cohorts from the first to the last it then proceeds to estimate the numbers of babies that will be born in each region. This operation is conveniently left until this stage so that all the relevant components that make up the population at risk of giving birth have been computed, and stored in the direct access file.

First, the alternative models for generating births are described together with the instructions for setting up the program to implement each. Then, the alternative methods for inputting births or fertility rates, and for selecting the populations at risk of giving birth are outlined.

3.16.2 The alternative ways of generating births

(i) A persons or unisex model

Births and associated fertility rates are classified by region of birth of infant and by age cohort of parent. The age cohorts considered are those of potentially fertile females within the age range 15 to 50.

Births we define, in slightly contracted accounts notation as

$K_a^{b(i)p}$

and fertility rates as

f_a^{ip}

where p refers to persons, the sex of those born, and a is an age cohort label indicating the age of the parents of those born.

Fertility rates are defined as

$$f_a^{ip} = K_a^{b(i)p} \;/\; \hat{K}_a^{Bip} \tag{3.103}$$

where the $\hat{K}_a^{Bip}$ variables are the populations of persons at risk of giving birth in region i while in cohort a. The equation that generates births from fertility rates is then

$$K_a^{b(i)p} = f^{ip} \quad \hat{K}_a^{Bip} \tag{3.104}$$

The age cohort range $a = a_1, \ldots, a_2$ is usually that of females, and this is how births are normally classified.

(ii) A female dominant model (sex independent)

Here we wish to estimate the numbers of male and female infants born. This is most conveniently accomplished by multiplying the births total by sex proportions, x^{im} and x^{if}

$$K^{b(i)m} = x^{im} \sum_{a=a_1}^{a_2} K_a^{b(i)p} \qquad (3.105)$$

where

$$x^{im} + x^{if} = 1 \qquad (3.106)$$

or $$x^{im} = 1 - x^{if} \qquad (3.107)$$

Fertility rates are defined as

$$f_a^{ip} = K_a^{b(i)p} / \hat{K}_a^{Bif} \qquad (3.108)$$

where the populations at risk consist of females only, and the births are generated by

$$K_a^{b(i)p} = f_a^{ip} \hat{K}_a^{Bif} \qquad (3.109)$$

A male dominant model would be one in which the populations at risk consisted only of males. This model requires that births be classified by age of father which is rarely done at regional level, and is a less accurate classification than that of age of mother (because the father's age may not be known for illegitimate births).

(iii) a female dominant model (sex dependent)

In the above model the sex of the infant was assumed to be independent of the age of mother. If the slight dependence of sex of child on age of parent was thought important this could be introduced into the fertility rates to give

$$K_a^{b(i)m} = f_a^{im} \hat{K}_a^{Bif} \qquad (3.110)$$

$$K_a^{b(i)f} = f_a^{if} \hat{K}_a^{Bif} \qquad (3.111)$$

if the requisite information was available.

(iv) A separate sex model

Occasionally it might be necessary to construct separate male and female accounts using the following equations

$$K_a^{b(i)m} = f_a^{im} \hat{K}_a^{Bim} \qquad (3.112)$$

and

$$K_a^{b(i)f} = f_a^{if} \hat{K}_a^{Bif} \qquad (3.113)$$

The age cohort range would still be that of females in the male equation.

How are these different ways of handling the sex structure of the population handled in the ABM program?

(i) A persons or unisex model

The nsex (number of sexes) parameter is set to 1 and the nfsex (index number of the fertile sex) is set to 1. Sex proportions should be input as 1.0's.

(ii) A female dominant model (sex independent)

The nsex parameter should be set to 2 or 3 (if persons' sums are required) and the nfsex parameter should be set to the index of the female sex (1 or 2 depending on the order of data assembly). Sex proportions for the first sex should be input in the data file. Those for the second sex are worked out as residuals from unity.

(iii) A female dominant model (sex dependent)

This can be approximated by running a separate sex model for females and a persons model for the two sexes combined, subtracting to yield male births.

(iv) Separate sex model

The parameters nsex and nfsex are set to 1, and two analyses, one for males and one for females, are carried out in sequence with sex proportions set to 1.0 in both cases.

3.16.3 Methods of inputting or estimating births

These are described in relation to the female dominant (sex independent) model.

(i) Births values input from the data file (jbirth = 1)

Values for births in the current period are read in from the data file for each of the fertile cohorts when the jbirth parameter is set to 1

$$K_a^{b(i)*(*)p}\ , \quad i = 1, \ldots, n; \quad a = a_1, \ldots, a_2$$

Fertility rates are computed using equation (3.108).

(ii) Fertility rates input from the data file (jbirth = 3)

Values for fertility rates in the current period are read in from the data file for each of the fertile cohorts when the jbirth parameter is assigned a value of 3

$$f_a^{ip}\ , \quad i = 1, \ldots, n; \quad a = a_1, \ldots, a_2$$

Births are computed using equation (3.109).

ITERATION NUMBER 67

A(I) VALUES

EAST ANG	SOUTH EA	REST BRT	REST WLD
0.999965	0.999963	0.999971	0.999967

B(J) VALUES

EAST ANG	SOUTH EA	REST BRT	REST WLD	TOTALS			
1.000035	1.000037	1.000030	1.000033	1.000035	1.000037	1.000030	1.000033

ACCOUNTS MATRIX VALUES

	EAST ANG 1	SOUTH EA 2	REST BRT 3	REST WLD 4	EAST ANG 5	SOUTH EA 6	REST BRT 7	REST WLD 8
EAST ANG	108147.	5727.	6959.	4260.	475.	12.	17.	9.
SOUTH EA	11208.	1250203.	59783.	82255.	23.	5489.	143.	170.
REST BRT	8783.	52189.	2896011.	86527.	18.	108.	14225.	207.
REST WLD	6868.	47924.	58658.	0.	14.	99.	140.	0.

BALANCING FACTORS ACHIEVED CONVERGENCEON ITERATION 67

REVISED ACCOUNTS MATRIX ADJUSTED TO CONSTRAINTS

	EAST ANG 1	SOUTH EA 2	REST BRT 3	REST WLD 4	EAST ANG 5	SOUTH EA 6	REST BRT 7	REST WLD 8	TOTALS 9
EAST ANG	108147.	5727.	6959.	4260.	475.	12.	17.	9.	125605.
SOUTH EA	11208.	1250203.	59783.	82255.	23.	5489.	143.	170.	1409274.
REST BRT	8783.	52189.	2896011.	86527.	18.	108.	14225.	207.	3058068.
REST WLD	6868.	47924.	58658.	0.	14.	99.	140.	0.	113704.
TOTALS	135006.	1356043.	3021411.	173042.	530.	5708.	14525.	386.	4706651.

ROW TOTALS OF INTERNAL ELEMENTS OF ADJUSTED ACCOUNTS

EAST ANG	SOUTH EA	REST BRT	REST WLD
125605.	1409273.	3058070.	113704.

COLUMN TOTALS OF INTERNAL ELEMENTS OF ADJUSTED ACCOUNTS

EAST ANG	SOUTH EA	REST BRT	REST WLD	TOTALS			
135006.	1356043.	3021411.	173042.	530.	5708.	14525.	386.

Figure 3.17 Details of the last iteration of the balancing factor model for the first cohort

(iii) Births from the previous period used (jbirth = 0)

Births for the current period are set equal to those of the previous period

$$K_{a}^{b(i)*(*)p}(p) = K_{a}^{b(i)*(*)p}(p-1) \qquad (3.114)$$

Fertility rates are computed using equation (3.108).

(iv) Fertility rates from the previous period used (jbirth = 2)

Fertility rates for the current period are set equal to those for the previous period

$$f_{a}^{ip}(p) = f_{a}^{ip}(p-1) \qquad (3.115)$$

Equation (3.109) is employed to compute births.

3.16.4 Populations at risk of giving birth

These are selected using the jparb parameter. They are defined below with respect to the female dominant (sex independent) model.

(i) Initial populations are used as populations at risk of giving birth if jparb is set to 1:

$$\hat{K}_{a}^{Bif} = K_{a}^{e(i)*(*)f} \qquad (3.116)$$

(ii) Average populations are used as populations at risk of giving birth if jparb is set to 2:

$$\hat{K}_{a}^{Bif} = \tfrac{1}{2}(K_{a}^{e(i)*(*)f} + K_{a}^{*(*)s(i)f}) \qquad (3.117)$$

(iii) Multiregional populations at risk are used as populations at risk of giving birth if jparb is set to 3:

$$\begin{aligned}\hat{K}_{a}^{Bif} = \; & (1)\ K_{a}^{e(i)s(i)f} + (0.5)\sum_{\substack{j\neq i\\ j\epsilon I}} K_{a}^{e(i)s(j)f} + (0.5)K_{a}^{e(i)s(R)f}\\ & + (0.5)\ K_{a}^{e(i)d(i)f} + (0.25)\sum_{\substack{j\neq i\\ j\epsilon I}} K_{a}^{e(i)d(j)f} + (0.25)K_{a}^{e(i)d(R)f}\\ & + (0.5)\sum_{\substack{j\neq i\\ j\epsilon I}} K_{a}^{e(j)s(i)f} + (0.5)K_{a}^{e(R)s(i)f}\\ & + (0.25)\sum_{\substack{j\neq i\\ j\epsilon I}} K_{a}^{e(j)d(i)f} + (0.25)K_{a}^{e(R)d(i)f}\end{aligned} \qquad (3.118)$$

Subroutines PARI, PARA and PARM are used to compute the three choices of population at risk.

3.16.5 Sex proportions

Only one set of sex proportions are used: one of
x^{im}, x^{if} or x^{ip}, $i=1, \ldots , n$
depending on the fertility model used.

3.16.6 Births model: outputs

Figure 3.18 shows the outputs of the births model: births and fertility rates for each fertile cohort, the sum of births over all fertile ages, the sex proportions adopted, and the sexed births that will be used in constructing the infant accounts. Note that Figure 3.18 displays only some of the fertile cohorts and that normally many more would be output.

3.17 Infant accounts

Once births for each region have been estimated and disaggregated by sex (if requested), the accounts based model subroutines are re-entered one further time to construct accounts for infants with the birth totals playing the role of initial populations. In assembling each of the components of the accounts, any of the options applicable to the existing population can be used; in addition there are other options, already described, that enable infant cohort components to be estimated or guestimated in the absence of the kind of statistics available for existing cohorts.

From a description of the methods and models underlying the menu of choices provided by the ABM program the paper turns to the detailed instructions for assembling the input data file in section 4.

```
INFANT ACCOUNTS     JBIRTH =  1   JPARB =  2   JSEXP =  1
--------------------------------------------------------------------------------
JBIRTH =  1  =  NEW VALUES ARE INPUT FOR BIRTHS IN THE CURRENT PERIOD
JPARB  =  2  =  THE AVERAGE POPULATIONS IN A COHORT ARE USED AS POPULATIONS AT RISK OF GIVING BIRTH
JSEXP  =  1  =  NEW VALUES FOR THE SEX PROPORTIONS ARE INPUT FROM THE DATA FILE
--------------------------------------------------------------------------------
BIRTHS AND FERTILITY RATES FOR FERTILE AGE COHORT   3  FOR PERIOD   1
--------------------------------------------------------------------------------
            EAST ANG   SOUTH EA   REST BRT
               3654.     33609.     90195.
            EAST ANG   SOUTH EA   REST BRT
            0.033263   0.030074   0.036085
TOTAL BIRTHS IN ALL REGIONS =     127458.
ALL REGION FERTILITY RATE =    0.034199
--------------------------------------------------------------------------------
BIRTHS AND FERTILITY RATES FOR FERTILE AGE COHORT   4  FOR PERIOD   1
--------------------------------------------------------------------------------
            EAST ANG   SOUTH EA   REST BRT
              32413.    312995.    742619.
            EAST ANG   SOUTH EA   REST BRT
            0.256912   0.234334   0.277743
TOTAL BIRTHS IN ALL REGIONS =    1088027.
ALL REGION FERTILITY RATE =    0.263098
--------------------------------------------------------------------------------
BIRTHS AND FERTILITY RATES FOR FERTILE AGE COHORT   5  FOR PERIOD   1
--------------------------------------------------------------------------------
            EAST ANG   SOUTH EA   REST BRT
              44900.    474663.   1021384.
            EAST ANG   SOUTH EA   REST BRT
            0.423054   0.404274   0.452264
TOTAL BIRTHS IN ALL REGIONS =    1540947.
ALL REGION FERTILITY RATE =    0.435465
--------------------------------------------------------------------------------
```

Figure 3.18 Births model: printout for the first three fertile cohorts

4. USERS' MANUAL

4.0 Introduction

In this section of the paper the instructions for assembly of the input data file are specified in a series of tables. These should be read in sequence from Table 4.0 through to Table 4.12 and the necessary input data assembled in a form suitable for input to the computer. Some users may simply wish to write down the necessary parameter choices and data items on sheets of paper and type these directly into the computer. Other users may want to prepare the data file on punched card coding sheeting and punch up a deck of cards for input to the computer.

Table 4.0 sets out the general layout of the data file. Something should be input for each item from (1) to (12) according to instructions given in Tables 4.1 through 4.12 for the first time interval being analyzed. Items (5) through (12) should be repeated for each subsequent period of analysis although for periods after the first there need not necessarily be anything to input.

The input data file should conclude with item (14) or recommence with item (1) if a second analysis is being carried out (not normally recommended).

Users may find the layout of the input data demanded in item (7) awkward from the point of view of the data sources they are likely to be using. The sequence will always be by age cohort. For example, assuming only 1 sex

component 1 for age cohort 1 (for sex 1)
component 2 for age cohort 1 (for sex 1)
: : : : : :
component Z for age cohort 1 (for sex 1)
: : : : : :
component 1 for age cohort 2 (for sex 1)
component 2 for age cohort 2 (for sex 1)
: : : : : :
component Z for age cohort 2 (for sex 1)
: : : : : :
component 1 for age cohort A (for sex 1)
component 2 for age cohort A (for sex 1)
: : : : : :
component Z for age cohort R (for sex 1)

where the components might be "initial populations", "interval surviving migrants", ... , "deaths". This data structure is a consequence of the way the program has been written. Users may wish to write their own data preparation routines to rewrite the "component x age cohort" organized source data into the "age cohort x component" form demanded by the ABM program.

Table 4.0 The general layout of an input data file for ABM

Item	Nature	Table in which details are given
(1)	title records (two)	Table 4.1
(2)	n-parameter record	Table 4.2
(3)	region names record(s)	Table 4.3
(4)	sex name record	Table 4.4
(5)	l-parameter record for period 1	Table 4.5
(6)	i-parameter record for "existing" age cohorts	Table 4.6 and associated subtables
(7)	data sets for "existing" age cohorts 1 (the first) to A (the last) as specified by item (6)	Table 4.7 and associated subtables
(8)	j-parameter record for births model	Table 4.8 and associated subtables
(9)	fertility data sets for fertile age cohorts α to β as specified by item (8)	Table 4.9
(10)	sex proportion records	Table 4.10
(11)	i-parameter record for the "infant" age cohort	Table 4.11
(12)	data set for the "infant" age cohort as specified by item (11)	Table 4.12
(13)	items or record sets (5) through (12) are repeated for as many periods as required	
(14)	concluding record containing 000 in columns 1 to 3	

4.1 Title Records

Table 4.1 Title records

Two 80 character records containing the title of the analysis being carried out. Any alphanumeric or punctuation characters allowed.

4.2 Analysis parameters (n-parameters)

Table 4.2 The n-parameter record for the parameters that control the dimensions of the population analysis for all periods

Parameter	Format	Columns	Description
nsex	i3	1-3	number of sexes (see note 1)
ninreg	i3	4-6	number of internal regions
ncohrt	i3	7-9	number of age cohorts
nalfa	i3	10-12	index number of first fertile age cohort
nbeta	i3	13-15	index number of last fertile age cohort
nfsex	i3	16-18	index number of fertile sex cohort
nform	i3	19-21	format switch (see note 2)
npriod	i3	22-24	number of periods of analysis
niter	i3	25-27	maximum number of iterations
nagein	i3	28-30	length of age interval (years)
tol	f9.8	31-39	tolerance level for convergence of accounts based model and constraints adjustment model

Note

1. Meaning of the nsex parameter values

nsex value	meaning	
1	any sex (males, females or persons):	nfsex must be 1
2	males and females:	nfsex can be 1 or 2
3	males, females and persons:	nfsex can be 1 or 2

If nsex equals 3, this simply instructs the program to add together male and female figures.

2. Meaning of the nform parameter values

nform	meaning
0	list directed input used
1	fixed format: fields of f10.0 or f10.6 used.

4.3 Region names

Table 4.3 Region names

One, two or three records are required in the following format

Variable	Format	Columns	Description
rnam 1(1) rnam 2(1)	a4 "	1-4 5-8	name of region 1
rnam 1(2) rnam 2(2)	" "	9-12 13-16	name of region 2
rnam 1(3) rnam 2(3)	" "	17-20 21-24	name of region 3
rnam 1(4) rnam 2(4)	" "	25-28 29-32	name of region 4
rnam 1(5) rnam 2(5)	" "	33-36 37-40	name of region 5
rnam 1(6) rnam 2(6)	" "	41-44 45-48	name of region 6
rnam 1(7) rnam 2(7)	" "	49-52 53-56	name of region 7
rnam 1(8) rnam 2(8)	" "	57-60 61-64	name of region 8
rnam 1(9) rnam 2(9)	" "	65-68 69-72	name of region 9
rnam 1(10) rnam 2(10)	" "	73-76 77-80	name of region 10

Notes

1. The number of region names required is that specified under the ninreg parameter on the n-parameter record plus two. Only as much of the record as is required need be filled in.

2. The ninreg + 1st name should be for "the rest of the world" or external zone.

3. The ninreg + 2nd name should be for labelling the total rows and columns

4. Example:- one record for 3 regions plus external zone plus table totals label.

```
        EAST ANGSOUTH EAREST BRTREST WLD  TOTALS
region name 1       2      3       4        5
```

4.4 Sex names

Table 4.4 Sex names

One record in the following format is required

Variable	Format	Columns	Description
snam 1(1)	a4	1-4	name of sex 1
snam 2(1)	"	5-8	
snam 1(2)	"	9-12	name of sex 2
snam 2(2)	"	13-16	
snam 1(3)	"	17-20	name of sex 3
snam 2(3)	"	21-24	

Notes

1. The number of sex names required is that specified under the nsex parameter on the n-parameter record. Only as much of the record as is required need be filled in.

2. Examples:-

PERSONS

sex name 1

MALE FEMALE PERSONS

sex name 1 2 3

4.5 Period parameters (l-parameters)

Table 4.5 The l-parameter record for the parameters that specify the years and tabulation levels associated with the population analysis of a period

Parameter	Format	Columns	Description	
lyear	i4	1-4	(Four digit) label for year in which period being analyzed starts	
lpriod	i4	5-8	length of period in years (see note 1)	
			value	meaning
ltab1	i4	9-12	1	age cohort by component tables only printed
			2	component by age cohort tables only printed
			3	both kinds of tables printed
			value	meaning
ltab2	i4	13-16	1	just populations printed after final iteration
			2	populations, components of growth and accounts printed after final iteration
			3	populations, components of growth, accounts and input variables printed after final iteration
			4	all components printed after final iteration
			5	all components printed at each iteration
			value	meaning
ltab3	i4	17-20	0	no printout of details of the constraints adjustment procedure
			≠0	printout requested of details of constraints adjustment procedure
			value	meaning
lauto	i4	21-24	0	previous period's values of the i-parameters used in current period
			1	new values of the i-parameters input for the current period

Notes

1. The length of the period(lpriod) should always be equal to the age interval (nagein)

4.6 Accounts based model parameters (i-parameters)

Table 4.6 The i-parameter record for the parameters that control the form of the accounts based model adopted in the current period

Parameter	Format	Columns	Component controlled	Description found in Table:
ipop	i3	1-3	Initial populations	Table 4.6.1
intmig	i3	4-6	Internal surviving migrants	Table 4.6.2
immig	i3	7-9	Surviving immigrants	Table 4.6.3
iparim	i3	10-12	Populations at risk of immigration	Table 4.6.4
iemig	i3	13-15	Surviving emigrants	Table 4.6.5
idth	i3	16-18	Deaths	Table 4.6.6
ipard	i3	19-21	Populations at risk of dying	Table 4.6.7
imin	i3	20-24	Minor flows (non-surviving migrants)	Table 4.6.8
instay	i3	25-27	Non-surviving stayers	Table 4.6.9
isstay	i3	28-30	Surviving stayers	Table 4.6.10
ifpop	i3	31-33	Final populations	Table 4.6.11
icon	i3	34-36	Constraints (marginal)	Table 4.6.12
iconk	i3	37-39	Constaints (matrix)	Table 4.6.13
iterbf	i4	40-43	Maximum number of iterations allowed in the constraints adjustment subroutine	
tolc	f10.8	44-53	Tolerance or error level adopted for the balancing factor procedure in the constraints adjustment subroutine	

Notes

1. Example for historical accounts (no constraints applied)

1 1 1 1 1 1 3 1 1 1 1 0 0 0 0.0

Table 4.6.1 The meaning of the ipop (initial population) parameter values

ipop value	meaning
0	The final populations of the previous period are used as the initial populations of the current period
1	New values are input for the initial populations
2	Births are used as initial populations in the infant age cohort
3	Initial populations are computed from the row accounting equations in a backcast version of the accounts based model
4	Initial populations are computed from equations involving mid-interval populations

Notes

1. If ipop equals 1 or 4, the necessary data items will need to be inserted in the appropriate place in the input data file.

Table 4.6.2 The meaning of the intmig (internal surviving migrants) parameter values

intmig value	meaning
0	The previous period's migrant flow values are used in the current period
1	New values are input for the internal surviving migrant flows
2	The previous period's migration rate values are used in the current period together with the initial populations as populations at risk
3	New values of the migration rates are input and used with the initial populations at risk
4	Infant migrant values are generated using half of the value of the parental migration rates multiplied by the birth's total to that cohort. See the text for a full explanation of the estimation model

Notes

1. If intmig equals 1 or 3, the necessary data items will need to be inserted in the appropriate place in the input data file.

Table 4.6.3 The meaning of the immig (surviving immigrants) parameter values

immig values	meaning
0	The previous period's migrant flow values are used in the current period
1	New values are input for the surviving immigrant flow
2	The previous period's immigration rate values are used in the current period together with the populations at risk specified by the iparim parameter. The iparim parameter should be the same value in the current and previous period
3	New values of the immigration rates are input and used with the populations at risk as specified by the iparim parameter. The populations at risk used in rate computation should match those specified in the iparim parameter
4	Infant migrants are estimated from the values for the initial cohort

Notes

1. If immig equals 1 or 3, the necessary data items must be included in the appropriate place in the input data file.

Table 4.6.4 The meaning of the iparim (population at risk for surviving immigrants) parameter values

iparim value	meaning
1	The initial populations are used as populations at risk
2	The total of immigrants (surviving and non-surviving) into the internal set of regions is used as the population at risk. This is only useful if such a flow is known exogeneously to the accounts based model
3	The final populations are used as populations at risk

Notes

1. The rates corresponding to these population at risk definitions may be described as follows:

iparim	corresponding rates
1	Pseudo-admission rates, but very useful. They can be net immigration rates.
2	Transition rates that distribute immigrants among internal final states.
3	Admission rates.

2. If iparim equals 2, the necessary total immigration flow must be read in from the data file.

Table 4.6.5 The meaning of the ienig (surviving emigrants) parameter values

ienig value	meaning
0	The previous period's surviving emigrant flow values are used in the current period
1	New values are input for the surviving emigrant flows
2	The previous period's emigration rate values are used together with the initial populations of the current period as populations at risk
3	New values for the emigration rates are input and used with the initial populations as populations at risk
4	Infant migrant values are generated using half of the value of the parental migration rates multiplied by the births totals to that parental cohort
5	Surviving emigrant values are estimated as residuals from the row accounting equation

Notes

1. If ienig equals 1 or 3, the necessary data items will need to be inserted in the appropriate place in the input data file.

Table 4.6.6 The meaning of the idth (total regional deaths) parameter values

idth values	meaning
0	The previous period's death flow values are used in the current period
1	New values are input for the deaths totals
2	The previous period's death rate values are used in the current period together with the populations at risk specified by the ipard parameter
3	New values are input for the death rates and these are used in conjunction with the populations at risk specified by the ipard parameter
4	Deaths are computed from the column accounting equation

Notes

1. If idth equals 1 or 3, the necessary data items must be inserted in the appropriate place in the input data file.

Table 4.6.7 The meaning of the ipard (population at risk of death) parameter values

ipard values	meaning
1	The initial populations are used as the populations at risk
2	The average populations in a cohort are used as the populations at risk
3	The multiregional populations at risk are computed

Table 4.6.8 The meaning of the imin (minor flow or non-surviving migrant) parameter values

imin values	meaning
1	Internal non-surviving migrants are estimated with the standard hypothesis: death rates of the region of destination are used
2	Internal non-surviving migrants are estimated with the death rates of the region of origin
3	Internal non-surviving migrants are estimated with the average of the origin and destination death rates

Note

Non-surviving emigrants are computed using the origin region death rates. Non-surviving immigrants are computed using the destination region death rates.

Table 4.6.9 The meaning of the instay (non-surviving stayers) parameter values

instay values	meaning
1	Non-surviving stayers are computed from the column accounting equation
2	Non-surviving stayers are computed from the row accounting equation
3	Non-surviving stayers are computed from a model equation

Table 4.6.10 The meaning of the isstay (surviving stayers) parameter values

isstay values	meaning
1	Surviving stayers are computed using the row accounting equation
2	Surviving stayers are computed using the column accounting equation

Table 4.6.11 The meaning of the ifpop (final population) parameter values

ifpop value	meaning
1	Final populations are computed from the column accounting equation
2	Final populations are input from the data file

Notes

1. If ifpop equals 2, the necessary data items must be inserted in the appropriate place in the input data file.

Table 4.6.12 The meaning of the icon (constraints:marginals) parameter values

icon value	meaning
0	No constraints are applied to the account matrices, which are thus unconstrained
1	Constraints on the account matrices are applied. The test for convergence is a relative test on balancing factors (see text)
2	Constraints on the accounts matrices are applied. The test for convergence is a relative test on each element of the accounts matrix (see text)
3	Constraints on the accounts matrices are applied. The test for convergence is an absolute test on each element of the accounts matrix (see text)

Table 4.6.13 The meaning of the iconk (constraints: matrix) parameter values

iconk value	meaning
0	No elements in the accounts matrix are held constant in the adjustment procedure
1	Some elements in the accounts matrix are held constant in the adjustment procedure. A matrix of 0's and 1's specifying which accounts elements are to be held constant (0's) and which can be adjusted (1's) must be input in the data file after the row and column constraints

4.7 Data sets for existing age cohorts

Table 4.7 Data sets for "existing" age cohorts

4.7.1 The sequence of data sets

Data sets are assembled for each age cohort and for each sex as required by the analysis parameters ncohrt (number of age cohorts) and nsex (number of sexes). The sequence is as follows:

data sets for age cohort 1, sex 1

data sets for age cohort 1, sex 2 (if required, that is, when nsex = 2 or 3)

data sets for age cohort 2, sex 1

data sets for age cohort 2, sex 2 (if required, that is, when nsex = 2 or 3)

⋮ ⋮ ⋮ ⋮ ⋮

data sets for age cohort R, sex 1

data sets for age cohort R, sex 2 (if required, that is, when nsex = 2 or 3)

where R is the number of age cohorts (ncohrt).

There will be:-

ncohrt such data sets if nsex = 1

2xncohrt such data sets if nsex = 2 or 3.

4.7.2 Structure of each data set

Each data set for an age cohort, sex group will consist of the following records.

(1) initial population record(s) (if required by ipop parameter setting)

(2) internal surviving migrant or migration rate records (if required by intmig setting)

(3) surviving immigrant or immigration rate record(s) (if required by immig parameter setting)

(4) total immigration record (if required by immig and iparin parameter settings)

(5) surviving emigrant or emigration rate record(s) (if required by iemig parameter setting)

(6) deaths or death rate record(s) (if required by idth parameter setting)

(7) final population record(s) (if required by ifpop parameter setting)

(8) row constraints record(s) (if required by icon parameter setting)

(9) column constraints records (if required by icon parameter setting)

4.7.3 Layout of record sets under list-directed input ("free format")

(The nform parameter on the n-parameter record is set to zero).

Vector record sets	Layout
(1) initial population record(s)	ninreg values, each separated from the next by a comma or blank or end of record
(3) surviving immigrant or immigration rate record(s)	
(5) surviving emigrant or emigration rate record(s)	
(6) deaths or death rate record(s)	
(7) final population record(s)	
(8) row constraints record(s)	ninreg values, each separated from the next by a comma or blank or end of record
(9) column constraints record(s)	2x(ninreg+1) values each separated from the next by a comma or blank or end of record
Matrix record sets	
(2) internal surviving migrant or migration rate records	ninreg setsof records each containing ninreg values, each separated from the next by a comma or blank. Dummy values (e.g. of zero) should be included for the diagonal or intra-region terms.
Scalar record sets	
(4) total immigration record	one value on one record.

Table 4.7 (Continued)

The number of records needed will depend on the length assigned by the user (80 characters is suggested, though a shorter record length may be more efficient of disk storage in certain circumstances) and the size of the numbers involved.

Example for one age cohort-sex group (ninreg = 3): East Anglia, South East and Rest of Britain accounts, persons, 1966-71, age cohort 0-4

125605 1409274 3058068	(A)
0 5739 6880	(B)
11171 0 58934	(B)
8869 52833 0	(B)
6900 48268 58283	(C)
4366 84055 89577	(D)
530 5708 14525	(E)

Notes

(A) initial population record (ipop = 1)

(B) internal surviving migrant records (intmig = 1)
Note the zero dummies for the diagonal elements.

(C) surviving immigrant record (immig = 1)

(D) surviving emigrant record (iemig = 1)

(E) deaths record (idth = 1)

4.7.4 Layout of record sets under fixed format input

The nform parameter on the n-parameter record is set to something other than zero (e.g.1).
All records are 80 characters in length.
The record format for flow variables is 8f10.0.
The record format for rate variables is 8f10.6.

Vector record sets	Layout
(1) initial population record(s)	n records of format 8f10.0 or 8f10.6 where n = smallest integer greater than or equal to (ninreg/8) evaluated as a real number
(3) surviving immigrant or immigration rate record(s)	
(5) surviving emigrant or emigration rate record(s)	
(6) deaths or death rate record(s)	
(7) final population record(s)	
(8) row constraints record(s)	n records of format 8f10.0 or 8f10.6 where n = smallest integer greater than or equal to (ninreg+1)/8 evaluated as a real number
(9) column constraints records	n records of format 8f10.0 or 8f10.6 where n = smallest integer greater than or equal to (2x(ninreg+1))/8 evaluated as a real number.

Matrix record sets	Layout
(2) internal surviving migrant or migration rate records	ninreg sets of n records of format 8f10.0 or 8f10.6 where n = smallest integer greater than or equal to (ninreg/8) evaluated as a real number
Scalar record sets	
(4) total immigration record	one record with one value in format f10.0 or f10.6

Example for one age cohort-sex group (ninreg=3): East Anglia, South East and Rest of Britain accounts, persons, 1966-71, age cohort 0-4.

125605	1409274	3058068	(A)
0	5739	6880	(B)
11171	0	58934	(B)
8869	52833	0	(B)
6900	48268	58283	(C)
4366	84055	89577	(D)
530	5708	14525	(E)

(A) initial population record (ipop=1)
(B) internal surviving migrant records (intmig=1)
(C) surviving immigrant record (immig=1)
(D) surviving emigrant record (iemig=1)
(E) deaths record (idth=1)

The ordering of items in each record set in the examples above

(1) initial population record	population of region 1	population of region 2	population of region 3
(2) internal surviving migrant records	stayers in region 1	migrants from region 1 to region 2	migrants from region 1 to region 3
	migrants from region 2 to region 1	stayers in region 2	migrants from region 2 to region 3
	migrants from region 3 to region 1	migrants from region 3 to region 1	stayers in region 3
(3) surviving immigrants record	immigrants to region 1	immigrants to region 2	immigrants to region 3
(6) deaths record	deaths in region 1	deaths in region 2	deaths in region 3

4.8 Births model parameters (j-parameters)

Table 4.8 The j-parameter record for the fertility model

Parameter	Format	Columns	Description
jbirth	i3	1-3	The birth model parameter (see Table 8.8.1)
jparb	i3	4-6	The population at risk of giving birth parameter (see Table 8.8.2)
jsexp	i3	7-9	The sex proportions parameter (see Table 8.8.3)

Table 4.8.1 The meaning of the jbirth (births) parameter values

jbirth value	meaning
0	The previous period's birth flows are used in the current period
1	New values are input for the births in the current period
2	The previous period's birth rates (fertility rates) are used in the current period together with the populations at risk specified by the jparb parameter
3	New values of the birth rates (fertility rates) are input and used in the current period together with the populations at risk specified by the jparb parameter

Notes

1. If jbirth equals 1 or 3, the necessary data items will need to be included in the input data file at the appropriate place.

Table 4.8.2 The meaning of the jparb (population at risk of giving birth) parameter values

jparb value	meaning
1	Initial populations are used as populations at risk of giving birth
2	The average populations in a cohort are used as populations at risk of giving birth
3	The multiregional populations at risk are used as populations at risk of giving birth

Notes

1. Populations at risk of giving birth are computed only for those cohorts specified as fertile by the NALFA and NBETA parameters.

Table 4.8.3 The meaning of the jsexp (sex proportions) parameter values

jsexp value	meaning
0	The previous period's sex proportion values are used in the current period
1	New values for the sex proportions are input from the data file

Notes

1. If jsexp equals 1, the necessary data items must be included in the input data file at the appropriate place.

4.9 Births data sets for fertile age cohorts

Table 4.9 Fertility data sets for fertile age cohorts α to β as specified by jbirth parameter

4.9.1 The sequence of data sets

Data sets are assembled for the fertile age cohorts as follows:

data set for age cohort α, the first fertile cohort

data set for age cohort $\alpha+1$, the second fertile cohort

⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮

data set for age cohort β, the last fertile cohort

There will be nfert = nbeta − nalfa +1 such cohorts where nalfa is the serial number of the first cohort and nbeta the serial number of the last cohort.

4.9.2 Structure of each data set

Each data set for a fertile cohort will consist of:

Either birth record(s)

or fertility rate record(s)

as required by the jbirth parameter.

4.9.3 Layout of record(s) under list directed input (the nform parameter or the n-parameter card is set to zero)

birth records	ninreg births values each separated from the next by a comma or blank or end of record
fertility records	ninreg fertility rate values each separated by a comma or blank or end of record.

Table 4.9 (Continued)

The number of records needed will depend on the length of record assigned by the user (80 characters is suggested) and the size of the numbers involved.

Example for age cohorts 3 to 10 (ninreg=3):
East Anglia, South East and Rest of Britain accounts, 1966-71

3654 33609 90195	(A)
32413 312995 742619	(B)
44900 474663 1021384	(C)
26411 312009 642932	(D)
11848 148999 316932	(E)
4623 58520 134069	(F)
876 10744 26305	(G)
45 509 1200	(H)

Notes

(A) Births to age cohort 10-14 to 15-19 in 1966-71
(B) " " " " 15-19 to 20-24 " "
(C) " " " " 20-24 to 25-29 " "
(D) " " " " 25-29 to 30-34 " "
(E) " " " " 30-34 to 35-39 " "
(F) " " " " 35-39 to 40-44 " "
(G) " " " " 40-44 to 45-49 " "
(H) " " " " 45-49 to 50-54 " "

4.9.4 Layout of record(s) under fixed format

The nform parameter on the n-parameter record is set to something other than zero (e.g.1). All records are 80 characters in length.

births records	nform sets of n records of format 8f10.0 where n = smallest integer greater than or equal to (ninreg/8) evaluated as a real number
fertility rate records	nform sets of n records of format 8f10.6 where n = smallest integer greater than or equal to (ninreg/8) evaluated as a real number

Example for age cohorts 3 to 10 (ninreg=3)
East Anglia, South East and Rest of Britain accounts, 1966-71

3654	33609	90195	(A)
32413	312995	742619	(B)
44900	474663	1021384	(C)
26411	312009	642932	(D)
11848	148999	316932	(E)
4623	58520	134069	(F)
876	10744	26305	(G)
45	509	1200	(H)

Notes

(A)-(H) as above under 8.9.3.

4.10 Sex proportions

Table 4.10 Sex proportion records

4.10.1 The assumptions are made that the sex of the infant at birth is not dependent on age of mother, and that the proportions in a second sex (if one is used in the analysis) are simply the complements to unity of those in the first. That is, the proportions of infants in the second sex are found by subtracting from 1 the proportions of the first sex.

4.10.2 Sex proportion record(s) under list directed input ("free" format)

These consist of ninreg sex proportion values each separated from the next by a comma or blank or end of record on as many records as are needed.

Values will generally range between .47 and .53, except when only one sex, persons, is being analyzed when they should be set to 1.0.

Example for East Anglia, South East and Rest of Britain accounts, 1966-71

.516 .514 .516

4.10.3 Sex proportion record(s) under fixed format

Ninreg sex proportions on n records of format 8f10.6, where n is the smallest integer greater than or equal to (ninreg/8) evaluated as a real number.

Example for East Anglia, South East and Rest of Britain accounts, 1966-71

.516 .514 .516

4.11 Accounts based model parameters for infant age cohort (i-parameters)

Table 4.11 The i-parameter record for the "infant" age cohort

The i-parameter record for the "infant" age cohort follows exactly the same format as the "existing" age cohorts i-parameter given in Table 4.6. The parameter choices are specified in the subtables associated with Table 4.6 The choices selected for the accounts components are likely to be different from those selected for the "existing" age cohorts.

4.12 Data sets for the infant age cohort

Table 4.12 The data set for the "infant" age cohort as specified by the i-parameter record preceding

The data set for the "infant" age cohort follows exactly the same layout as specified for the "existing" age cohorts in Table 4.7 and associated subtables under the control of the parameters specified in the i-parameter record for infants. The numbers, however, are likely to be smaller in size because of the unique nature of the infant cohort.

4.13 Repetition for further periods

Items (5) through (12) can be repeated for as many periods as desired.

4.14 Finishing off

A record of three zeroes as the first three characters should be appended to the data file to conclude the analysis.

5. EXAMPLE OF A DATA FILE

Figures 5.1 and 5.2 show the beginning and end of a typical data file for the three region systems of East Anglia, the South East and Rest of Britain, with the population disaggregated into 16 age cohorts.

5.1 Figure 5.1

The first two records of type A: they contain titles for the analysis.

The third record of type B consists of the analysis parameters and indicates, reading from right to left, that the number of sexes in the population is 1, the number of regions 3, the number of age cohorts 16, the 3rd of which is the first fertile cohort, the 10th of which is the last fertile cohort. The next parameter value, 1, indicates that the first (and only) sex is the fertile one. The following 0 instructs the program to read the data for each cohort as a list with items separated by spaces, commas or end of records. The number of priods of analysis is 2, the number of iterations of the accounts based model allowed is 4, the age interval is 5 years and the tolerance level for testing for convergence of the accounts based model is .0001.

The fourth record, C, contains the region names and the fifth record, D, the sex name.

Record E, the next, holds the period parameters: namely the year in which the first analysis period starts, 1966, the length of the period, 5 years; the next three numbers give the settings for the tabulation parameters. The 1 means that only age cohort by component tables are printed; the 2 means that populations, components of growth and accounts are printed for each age cohort (but only once after the final iteration); the 0 indicates that no tabulation of the balancing factor-constraints procedure is to be printed since it has not been requested. The final parameter, set to 1, indicates that the program should look for and read in a set of accounts based model parameters for this period on the next record.

The next record of type F contains the model parameter settings. Most are set to 1 indicating that the model is designed to produce 'standard' historical unconstrained accounts in which most components are read in as data. The line 3 refers to the population at risk of dying which is selected to be the multiregional one, and the three final zeroes mean that the user is not applying any constraints to the accounts.

There then follow blocks of records of type G, each block

```
Record
 type

A        A TEST DATA SET FOR ABM (FROM WP 229)
A        EAST ANGLIA AND SOUTH EAST REGION ACCOUNTS, 1966-71, PERSONS
B          1  5 16  5 10  1  0  2  4  5  .0001
C        EAST ANGSOUTH EAREST BRTREST OLD  TOTALS
D         PERSONS
E        1966   5   1   4   0   1
F          1  1  1  1  1  1  5  1  1  1  1  0  0 .000
G a      125005 1409274 3058068
  b      0 5739 6880
  b      11171 0 58934                           data for
  b      6869 52833 0                            age cohort 1
  c      6900 48268 58283
  d      4566 84055 39577
  e      550 5708 14525
G a      112901 1238011 2733596
  b      0 3972 4578                             data for
  b      7513 0 39795                            age cohort 2
  b      6219 37861 0
  c      3840 45949 55146
  d      4011 77256 63011
  e      787 1786 4541
G a      107773 1119789 2520237
  b      0 3893 3673                             data for
  b      5710 0 31736                            age cohort 3
  b      5814 40077 0
  c      3800 49389 40969
  d      3334 63670 68014
  e      272 2510 5667
```

Figure 5.1 <u>Part one of an example data file for East Anglia, the South East and Rest of Britain accounts</u>

<u>Record type</u>

- A – titles
- B – analysis parameters (n-set)
- C – region name
- D – sex name
- E – period parameters (i-set)
- F – model parameters (i-set)
- G – data set for an age cohort
  - a – initial populations
  - b – surviving internal migrants
  - c – surviving immigrants
  - d – surviving emigrants
  - e – deaths

pertaining to one age cohort, all of the same structure. Each block contains seven records. The first, type a, holds the initial populations of the three regions. The next three, type b, hold the surviving internal migrant numbers in a 3 x 3 matrix with zero diagonals. The next record, type c, contains the surviving immigrant flows, followed by a type d record containing surviving emigrant record and a type e record containing deaths.

5.2 Figure 5.2

The G blocks of records then repeat until the end of the data file is reached, as shown in Figure 5.2.

This contains the last G block for the 16th cohort, followed by a record of type H containing the births model parameters. The first parameter, set to 1, indicates that births values are to be read in; the second, with a value of 2, that the average population is to be used as the population at risk of giving birth, and the third, set to 1, instructs the program to look for and read in the sex proportion values.

A block of records, type I, then follow containing the births in each region, one record per fertile cohort. This block is followed by a record, type J, containing the sex proportions for infants, and another parameter record, type F again, containing the accounts based model parameters for the infant cohort.

Then follows the block of data for the infant cohort. This differs from the previous data blocks in having no initial populations. Births take their place.

With the E type record that follows the infant cohort data, the instructions for a new period of analysis begin. These instructions consist of an E type record of period parameters, and F type record of model parameters for existing age cohorts, and H type record of births model parameters and an F type record of infant cohort model parameters. No further data is input for the second period as the parameter settings instruct the program to use data from the first period, stored on the direct access desk file, in various ways.

```
Record
type

G a          80768 801795 1458635
  b          0 549 754
  b          2125 0 6384
  b          1729 6083 0
  c          81 1254 1421
  d          86 1782 1501
  e          37600 309517 710218
H              1  2  1
I 3          5654 33609 90195
  4          32415 312995 742019
  5          44900 474665 1021564
  6          26411 312009 642932
  7          11848 148999 316932
  8          4625 58520 134069
  9          876 10744 20505
 10          45 509 1200
J            1. 1. 1.
F              2   1   1   1   1   1   5   1   1   1   1   0   0 .000
G b          0 4201 3658
  b          6519 0 53215
  b          8889 30178 0
  c          3420 24610 31471
  d          2350 44527 40972
  e          2047 24059 55207
E            1971   5   2   1   0   1
F             0   2   0   1   0   2   5   1   1   1   1   0   0 .000
H             2   2   0
F             2   2   0   1   0   2   5   1   1   1   1   0   0 .000
B            000
```

Figure 5.2 Part two of an example data file for East Anglia, the South East and the Rest of Britain accounts

Record type

H - births model parameters (j-set)
I - data set for births for fertile age cohorts
J - sex proportions
K - model parameters for infants (i-set)

6. RUNNING ABM ON THE AMDAHL VM470

The instructions for running the program are here specified for the Amdahl VM470 at the University of Leeds. They should, however, be implementable also on IBM machines running under the CMS system. Other users will have to adapt these instructions to their own machine's operating system.

6.1 The direct access disk file

In order to run the ABM program the user must have enough space on his or her disk to accommodate the direct access disk file that will be generated by his or her analysis. The number of records that this file will contain must be specified in the program source version and the program recompiled.

The direct access disk file is used to store the accounts flow and rate-variable associated with the regional system being analyzed for one period. The number of records needed is a function of

the number of internal regions	N
the number of age cohorts	A
the number of sexes	X
the number of fertile age cohorts	F

Namely, the

$$\begin{aligned} \text{number of records required} = \; & 2\ (A+3)X(7M+NM+4) \\ & + (M+1)(F+1)+(M+1)X \\ & + (M+1)(F+1)+M \\ & + AXM \end{aligned} \qquad (6.1)$$

where M = the smallest integer equal to or greater than (n/8)

For example

if $N = 3$

$A = 4$

$X = 1$

$F = 2$

then $M = 1$

number of records $= 238 + 8 + 7 + 4 = 257$

A number equal to or greater than 257 must then be entered in the define file statement in the program and in the file definition statement associated with data set 3 when the program is executed.

Rather than make a mistake with the arithmetic in evaluating equation (6.1) users may wish to run the program DANREC listed in Appendix A.4

6.2 Editing and compiling the program

6.2.1 When it is not necessary

If the ABM program exists in a compiled form with a define file statement containing a large enough number of records, then editing and compilation of the program will not be necessary. The TEXT version of the program that resides in the School of Geography program library can be used.

Having logged on the computer you should type the following exec command

abm fn1 fn2 nrec

where fn1 is the filename of the input data file, fn2 is the filename of the direct access disk file and nrec is the number of records in that direct access disk file. You have access to ABM TEXT on the library b disk and are running an analysis using an exec similar to Example 2 in section 6.4.

6.2.2 When it is necessary

If the number of records specified in the define file statement is insufficient, the program needs editing and recompilation. Users unfamiliar with such procedures or with restricted disk space should request the necessary modifications by the author. For those users with their own copy of the source version of the program, the instructions are as follows:

(1) Get access to the program

logon userid password

files

edit abm fortran

locate/define file 3(

(2) The editor at this point will type out the contents of the record in the program containing "define file 3(".

Examine the record which will be of the form

define file 3(x,80,e,i1)

where x will be an integer number specifying the number of records in the direct access file.

(3) Then

change /x/y/

where y is the number of records desired in the direct access file.

(4) Preserve the edit by issuing the instruction

file

(5) Then compile the program

forthxe abm

6.3 File definition statement

In order to run the program you will need to provide the operating system of the computer with definitions of the files being used in file definition statements.

The program makes the following assumptions:

(1) the information assembled by the user is read into the program from data set 1;

(2) the results of the analysis are written out to data set 2;

(3) the direct access disk file is addressed for reading from and writing to as data set 3.

There are many ways in which the file definition statements could be specified depending on the analysis being carried out by the user and the disk space available to the user.

Example 1

```
fi 1 disk pop data  (recfm fb lrecl 80 blksize 400
fi 2 disk pop results  (recfm fb lrecl 132 blksize 660
fi 3 disk popda data  (xtent 300 recfm fb lrecl 80 blksize 400
```

Here the input data is on the disk in the filename 'pop' with the filetype 'data', a file of fixed length, blocked records of length 80 characters in blocks of 400 characters. The program output is written to a disk file called 'pop' with the filetype 'results'. This is a fixed length, blocked record file with records of 132 characters in blocks of 660. Finally, the direct access file is called 'popda' with filetype 'data' with fixed record length of 80 characters in blocks of 400. The number of records in this file must be specified by the xtent option: in this case 300 records are assumed.

Example 1's file definitions could be very expensive of disk space so that for a problem involving a lot of output we would want to change the definitions as follows:

Example 2

```
fi 1 disk pop data  (recfm fb lrecl 80 blksize 400
fi 2 printer  (recfm fb lrecl 132 blksize 660
fi 3 disk popda data  (xtent 300 recfm fb lrecl 80 blksize 400
```

The output is now sent directly to the printer rather than the user's disk space.

However, for problems involving single year age cohorts, the user may need to save even more diskspace.

Example 3 shows how this might be done.

Example 3

```
fi 1 reader  (recfm fb lrecl 80 blksize 400
fi 2 printer  (recfm fb lrecl 132 blksize 660
fi 3 disk popda data  (xtent 300 recfm fb lrecl 40 blksize 400
```

In this example the input data file is located in the user's reader, the output is sent to his or her printer. The direct access file must remain on the disk, but the length of the record has been reduced to 40 characters because only 4 internal regions were involved in the analysis, thus saving some disk space. If the number of regions, in the analysis is less than 8, disk space may be saved by specifying the record length as n x 10 with a suitable adjustment of the blocksize parameter.

6.4 Running the program

To run the program issue the following commands:

```
setup fortran
fi 1 etc. )
fi 2 etc. )  as selected following the guidelines in
fi 3 etc. )     section 6.4
load abm
start (clear
```

The first command gives the user access to the necessary fortran libraries; the second set of commands are the file definition statements needed; the text version of the program is then loaded; and execution commences with the start command. The (clear option ensures all core store assigned to which program variables are assigned is initially cleared (that is, set to zero).

If many runs of the program are to be made the user may wish to set up his own exec procedure to do all of the above, varying each time the names of the data files being used.

Example 1

A simple file called abm exec might look like this:

```
setup fortran
fi 1 disk pop data  (recfm fb lrecl 80 blksize 400
fi 2 printer  (recfm fb lrecl 132 blksize 660
fi 3 disk popda data  (xtent 300 recfm fb lrecl 80 blksize 400
start (clear
```

Execution would be initiated by typing the command abm.

Each time the problem was run the file 'abm exec' would need to be edited and the filenames changed.

Example 2

A more convenient abm exec file would involve general arguments replacing the filenames and xtent option number.

```
setup fortran
fi 1 disk &1  data  (recfm fb lrecl 80 blksize 400
fi 2 printer  (recfm fb lrecl 132 blksize 660
fi 3 disk &2  data  (xtent & 3 recfm fb lrecl 80 blksize 400
load abm
start (clear
```

Execution would be initiated by typing, for example,

```
abm pop popda 300
```

6.5 Disk space requirements

6.5.1 Minimization of disk space requirements.

The minimum disk space required by the user for running large problems is the space occupied by the direct access data file, assuming that the text version of the program is accessible on another user's disk space (for example, the geography program library's disk space).

6.5.2 Maximum problem size.

The maximum size of problem that the program is designed to handle involves

20 internal regions

95 age cohorts+

3 sexes

35 fertile age cohorts

In this case M (magic1) is 3 and

the number of records required = 86418

This is probably well in excess of the space available to the user even by special request. Considerable savings could be effected by a slight aggregation of the system to

16 internal regions

95 age cohorts

2 sexes (just the summation of sexes is foregone)

35 fertile age cohorts

In this case M (magic1) is 2 and

the number of records required = 32748

+ There is no explicit limit to the number of age cohorts but users are unlikely to want to go beyond age 95, with the last cohort being persons aged 95 and over.

It should prove possible to modify the program so that problems are handled in two blocks the first involving all ages up to and including the last fertile age cohort, and the second involving age cohorts beyond the fertile. The system would then consist of

16 internal regions

50 age cohorts

2 sexes

35 fertile age cohorts

In this case the number of records required on the direct access disk file is 17808. This number should be available on an 8 cylinder disk.

A final way to reduce the amount of diskspace required would be to run the two sexes separately adopting a male dominant birth model for males, or together, reducing the number of sexes to one. The system would then consist of

16 internal regions

50 age cohorts

1 sex

35 fertile age cohorts

and the number of records required would be 9016.

6.5.3 Being realistic about the level of disaggregation.

In practice, reliable estimation of the variables involved in a multiregional population model with both many regions and many age cohorts will be difficult. If age detail is required then fewer regions should be employed simultaneously in the multiregional system. If more detail about interactions between regions is required age detail will need to be sacrificed.

7. CONCLUDING REMARKS

The ABM program provides the population analyst with a modelling system designed to make possible the compilation of historical, estimated and projected population accounts and their associated population stocks and components of growth on an age cohort and sex disaggregated basis.

Further development of the program is intended, but the current version is now made generally available in order that further testing with real world problems can be carried out. The author would be grateful if users experiencing problems in running the program would communicate with him in order that the program can be further improved.

Further program developments are planned. Among these will be the following:

(i) A further reduction in the direct access disk space requirements will be sought. It should prove possible to reduce these by factor of 4.

(ii) Routines for substituting a few model parameters for extensive vectors or matrices of rates will be introduced to make the task of projection easier.

The Appendices which follow are intended as aids to the implementation of the ABM program by a programmer. Appendix A.1 contains a brief guide to the programmer including a flow chart of the sequence of subroutine calls in the main segment, and details of the ordering and layout of the direct access file. Appendix A.2 consists of instructions for obtaining a full listing of the program. Appendix A.3 contains details of how to obtain listings of an input data file and output results file for checking purposes. Appendix A.4 gives the text of the DANREC program (direct access number of records) for computing the size of the direct access data file before running ABM.

References

Castro, L. (1980) Aggregated data transformation: transitions and moves. Draft of unpublished paper (personal communication).

Chiang, C.L. (1968) An introduction to stochastic processes in biostatistics. Wiley, New York.

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

Jenkins, J.C. and Rees, P.H. (1977) Computer programs for aggregate accounts-based and associated forecasting models of population. Working Paper 205/Computer Manual 2, School of Geography, University of Leeds.

Jenkins, J.C. and Rees, P.H. (1978) A manual for using multipopl. A program for calculating aggregate historical accounts based population models, with the option of a simple growth matrix forecasting model. Computer Manual 9, School of Geography, University of Leeds.

Ledent, J. (1980) Multistate life tables: movement versus transition perspectives. Environment and Planning A, 12, 533-562.

Ledent, J. and Courgeau, D. (1980) Migration and settlement: France. Research Report, International Institute for Applied Systems Analysis, Laxenburg, Austria, forthcoming.

Ogilvy, A. (1980a) IInter-regional migration since 1971: an appraisal of data from the National Health Service Central Register and Labour Force Surveys. Occasional Paper 20, Office of Population Censuses and Surveys, H.M.S.O., London.

Ogilvy, A. (1980b) Migration - the influence of economic change. Futures, 11, 5, 383-394.

Plessis-Fraissard, M. and Rees, P.H. (1977a) A computer program for constructing age-sex multiregional population accounts (DAME): a full description. Working Paper 169, School of Geography, University of Leeds.

Plessis-Fraissard, M. and Rees, P.H. (1977b) A computer program for constructing age-sex multiregional population accounts (DAME): a users' manual. Working Paper 170, School of Geography, University of Leeds.

Plessis-Fraissard, M. and Rees, P.H. (1978) Computer program 1: a program for constructing age-sex disaggregated multiregional accounts (DAME). Environment and Planning A, 10, 225-226.

Rees, P.H. (1977) The measurement of migration from census data and other sources. Environment and Planning A, 10, 705-726.

Rees, P.H. (1978) Problems of multiregional population analysis: data collection and demographic accounting. Working Paper 229, School of Geography, University of Leeds, published in part as Rees, P.H. (1979) same title, Sistemi Urbani, 1, 3, 3-32.

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

References (continued)

Rees, P.H. (1980) Multistate demographic accounts: measurement and estimation procedures. Environment and Planning A, 12, 499-531.

Rees, P.H., Smith, A.P. and King, J.R. (1977) Population models. Chapter 3 in A.G. Wilson, P.H. Rees and Leigh, C.M. (ed.) Models of cities and regions. Wiley, Chichester, Sussex.

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

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

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

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

Stillwell, J.C.H. (1980) Theoretical and empirical analysis of internal migration: a bibliography. Working Paper 284, School of Geography, University of Leeds.

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

Stone, R. (1975) Towards a system of a social and demographic statistics. United Nations, New York.

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

APPENDICES

A.1 A BRIEF GUIDE TO ABM FOR PROGRAMMERS

This Appendix is designed as an aid to programmers attempting to implement the ABM program on a new machine. The guide is fairly brief as comment statements have been inserted at frequent intervals in the program text. The sequence and functions of the program subroutines are explained. And the layout of the direct access file is given in some detail, as this is difficult to reconstruct from the program text.

A.1.1 The sequence of subroutines

MAIN

The main segment reads in the job titles, the analysis parameters (n-set), the region names, the sex names, the period parameters (l-set), the accounts based model parameters for existing age cohorts (i-set) and the infant cohort (another i-set). The main segment calls the sequence of accounts based model subroutines BEGPOP ... BIRTHS described in section 3 of the paper and two "reporting" subroutines AGESUM, which computes age aggregated totals, and AGETAB, which prints out the results of the analysis in age classified tables, if the user so wishes.

There follows a set of utility subroutines.

MAGIC

This subroutine computes a number of integer quantities (magic 1 ... magic 9) needed for accessing the direct access file in most other subroutines.

ITELL

This reports in detail on the meaning of the i-parameter choices selected by the user, so that the model underlying the results of any run of the program is clearly identified in words.

VECOUT

This subroutine prints out a vector of regional accounts variables when requested within other subroutines.

MATOUT

This subroutine prints out a matrix of regional accounts variables when requested within other subroutines.

DAFSCA

This subroutine reads from or writes to the direct access file the value of a scalar accounts variable.

DAFVEC

This subroutine reads from or writes to the direct access file the values of a vector accounts variable.

DAFMAT

This subroutine reads from or writes to the direct access file the values of a matrix accounts variable.

The next set of subroutines deal with the components of the accounts, one by one. The methods and models used in each have been discussed extensively in section 3.

Subroutine	Component read in or computed
BEGPOP	Initial populations
MIG11	Surviving internal migrants
MIG01	Surviving immigrants
MIG10	Surviving emigrants
DEATHS	Deaths
PARI	Populations at risk: initial
PARA	Populations at risk: average
PARM	Populations at risk: multiregional
	The population at risk subroutines are called from within other accounts based model subroutines.
MINFLO	Non-surviving internal migrants, immigrants and emigrants
NSSTAY	Non-surviving stayers
SSTAY	Surviving stayers
ENDPOP	Final populations
TEMIG	Emigrant totals
TIMMIG	Immigrant total

Then follows a set of subroutines that do a variety of different jobs connected with the accounts based model.

ACCASS

This subroutine assembles accounts variables into one matrix.

ACCNTS

This subroutine arranges the printout of whole accounts tables.

VERGE

This subroutine tests whether the accounts based model has converged.

PERSON

This sums male and female accounts variables if requested.

CONSTR

This subroutine reads in values for row and column n constraints and adjusts the accounts matrix to these new marginals.

REWRIT

This subroutine writes the values of the revised accounts variables to the direct access file afresh.

BIRTHS

The births sub-models are implemented through this subroutine (see section 3).

IBTELL

A detailed description of the meaning of the birth model parameters is provided by this subroutine.

KSISTO

This subroutine arranges the storage of the final populations of each age cohort on the direct access file in a stack for use in projection runs.

AGESUM

This subroutine produces summary accounts for all existing age cohorts and for all age cohorts (existing and infant).

COG

This subroutine arranges the computation and printout of components of growth tables.

A final set of subroutines arrange for the printout of accounts variables in age classified tables component by component.

AGETAB

This subroutine arranges for the printout of age classified tables of accounts variables, if requested. To do this the following four subroutines are called.

SCATAB

This prints out a scalar accounts variable in age classified form

VECTAB

This prints out a vector accounts variable in age classified form.

MATTAB

This prints out a matrix accounts variable in age classified form.

COGTAB

This prints out components of growth variables in age classified form.

A.1.2 The layout of the direct access file

The direct access file consists of five blocks of variables:

(i) the stocks and flows variables of the accounts (k variables);

(ii) the rates variables of the accounts (the h variables);

(iii) the births variables (K^b variables);

(iv) the fertility rates variables (the f variables);

(v) the final population stack (the $K_a^{*(*)s(i)}$ variables).

The number of records occupied by a particular variable or block of variables is defined in terms of a set of general integer numbers, magic1 to magic9. Their definitions are set out in Table A.1.

Table A.1 Definitions of the magic variables

Preliminary FORTRAN definitions

ninreg = number of internal regions (N)
ncohrt = number of age cohorts (A)
nsex = number of sexes (X)

All records are of length 80 characters

Magic variable definitions

magic 1 = ifix ((float/ninreg/8.) + .875)
= number of records needed to store a vector of variables values for N regions in format 8 f 10.0 or 8 f 10.6

magic 2 = ninreg * magic 1
= number of records needed to store a N by N matrix of variable values in format 8 f 10.0 or 8 f 10.6

magic 3 = 7 * magic 1 + 2 * magic 2 + 4
= number of records needed to store all accounts variables for one age cohort for one sex

magic 4 = nsex * magic 3
= number of records needed to store all accounts variables for one age cohort for all sexes

magic 5 = (ncohrt + 3) * magic 4
= number of records in the K variable or h variable block for all the existing age cohorts, plus the infant and two summary cohorts

magic 6 = magic 1 + 1
= number of records per age cohort in the births block of the direct access file

magic 7 = ((n fert + 1) * magic6) + (nsex * magic 6)
= number of records for all age cohorts in the births block of the direct access file

magic 8 = ((n fert + 1) * magic 6) + magic 1
= number of records for all age cohorts in the fertility rate block of the direct access file

magic 9 = ncohrt * nsex * magic 1
= number of records needed for the final population stack

Table A.2 gives the layout of the records stocks and flows variables of the accounts for one age cohort for one sex. There will be (nsex * (ncohrt+3)) of these units in the K variables block of the direct access file.

Table A.3 gives the layout of the records for rates variables of the accounts for one age cohort for one sex. There will be (nsex * (ncohrt+3)) of these units in the h variables block of the direct access file.

Table A.4 provides the layout of the births block of the direct access file; Table A.5 gives the slightly shorter layout of the fertility rates block; and finally Table A.6 gives the details of the final population stack.

Table A.2 The layout of the records for one age cohort for one sex in the K variables block of the direct access file

Accounts variable: component name	FORTRAN name	Number of records required for unit	Record number that starts unit of records: ((iage-1) * magic 4 + (msex - 1) magic 3) +
initial populations	ki	magic 1	1
total immigrants	koi	1	magic 1 + 1
deaths	kdi	magic 1	magic 1 + 2
total non-surv emigrants	kdo	1	2 * magic 1 + 2
surviving irternal migrants & stayers	ksii	magic 2	2 * magic 1 + 3
surviving immigrants	ksoi	magic 1	magic 2 + 2 * magic 1 + 3
surviving emigrants	ksio	magic 1	magic 2 + 3 * magic 1 + 3
non-surviving internal migrants & stayers	kdii	magic 2	magic 2 + 4 * magic 1 + 3
non-surviving immigrants	kdoi	magic 1	2 * magic 2 + 4 * magic 1 + 3
non-surviving emigrants	kdio	magic 1	2 * magic 2 + 5 * magic 1 + 3
final populations	ksi	magic 1	2 * magic 2 + 6 * magic 1 + 3
total surviving emigrants	kso	1	2 * magic 2 + 7 * magic 1 + 3
grand total all components	kt	1	2 * magic 2 + 7 * magic 1 + 4

Note

1. iage = index number of age cohort

2. msex = index number of sex

Table A.3 The layout of records for one age cohort for one sex in the h variable block of the direct access file

Accounts variable: name of rate	FORTRAN name	Number of records required for unit	Record number that starts unit of records: magic 5 + ((iage-1) * magic 4) + (msex -1) * magic 3) +
total transition rates	hi	magic 1	1
total immigration rates	hoi	1	magic 1 + 1
death rates	hdi	magic 1	magic 1 + 2
total non-surviving emigration rate	hdo	1	2 * magic 1 + 2
transition rates for surviving internal migrants and stayers	hsii	magic 2	2 * magic 1 + 3
immigration rates	hsoi	magic 1	magic 2 + 2 * magic 1 + 3
emigration rates	hsio	magic 1	magic 2 + 3 * magic 1 + 3
transition rates for non-surviving internal migrants and stayers	hdii	magic 2	magic 2 + 4 * magic 1 + 3
non-surviving immigration rates	hdoi	magic 1	2 * magic 2 + 4 * magic 1 + 3
non-surviving emigration rates	hdio	magic 1	2 * magic 2 + 6 * magic 1 + 3
population change rates	hsi	magic 1	2 * magic 2 + 6 * magic 1 + 3
total surviving emigration rate	hso	1	2 * magic 2 + 7 * magic 1 + 3
grand total	ht	1	2 * magic 2 + 7 * magic 1 + 4
all rates		magic 3	

Note

1. iage = index number of age cohort

2. msex = index number of sex

Table A.4 The layout of records for the births block of the direct access file

Age cohort nsex	Births variable Description	FORTRAN name	Number of records required for unit	Record number that starts unit of records: ((2 * magic 5)+(ifaged -1) * magic 6)) +
BIRTHS BY AGE COHORT				
first fertile cohort	births births total	bi bit	magic 1 1	1 magic 1 + 1
⋮				
last fertile cohort	births births total	bi bit	magic 1 1	1 magic 1 + 1
summary cohort	births sum births total sum	bi bit	magic 1 1	1 magic 1 + 1
BIRTHS BY SEX				+(nfert + 1)*magic 6) + (msex * magic 6) +
sex one	births summed over age for sex one total of above	kbi kbit	magic 1 1	1 magic 1 + 1
⋮	⋮	⋮	⋮	⋮
sex nsex	births summed over age for nsex	kbi kbit	magic 1 1	1 magic 1 + 1

Note

1. ifaged = index number of fertile age cohort

Table A.5 The layout of records for the fertility rates block of the direct access file

Age cohort or sex	Fertility rate variable	FORTRAN name	Number of records required for unit	Record number that starts unit of records: ((2 * magic 5) + (magic 7) + (ifage - 1 * magic 6))
FERTILITY RATES BY AGE COHORT				
first fertile cohort	fertility rate	fri	magic 1	1
	fertility rate	frit	1	magic 1 + 1
⋮				
last fertile cohort	fertility rate	fri	magic 1	1
	system fertility rate	frit	1	magic 1 + 1
summary cohort	total fertility rate	tfri	magic 1	1
	system total fertility rate	tfrit	1	magic 1 + 1
	sex proportions	sexpi	magic 1	((nfert + 1) * magic 6) + 1

Table A.6 The layout of records in the final population stack block of the direct access file for one age cohort for one sex

Accounts variable: component name	FORTRAN name	Number of records required for unit	Record number that starts unit of records: (2 * magic 5 + magic 7 + magic 8) + ((iage - 1)* magic 1) +
final populations	ksi	magic 1	1

Note

1. There will be nsex * ncohrt * magic 1 records in final population stack

A.2 A LISTING OF THE PROGRAM, A TEST DATA FILE AND TEST RESULTS

Since the listing involves more than 75 pages in its shortest form, it is not given here. The author would be happy to provide listings of the program together with machine readable code on magnetic tape at cost to interested users. Users can also be povided with a test data file and test results with which to check their version of the program.

A.3 A LISTING OF THE DANREC PROGRAM

```
FILE: DANREC   FORTRAN  A   LEEDS UNIVERSITY VM/BSE


      COMMON /A1/ NINREG,NCOHRT,NSEX,NFERT
      COMMON /M/ MAGIC1,MAGIC2,MAGIC3,MAGIC4,MAGIC5,
     *MAGIC6,MAGIC7,MAGIC8,MAGIC9
      WRITE (2,100)
  100 FORMAT(1H ,'PLEASE INPUT THE NUMBER OF REGIONS IN THE ANALYSIS.')
      READ(5,*) NINREG
      WRITE(2,101)
  101 FORMAT(1H ,'PLEASE INPUT THE NUMBER OF AGE COHORTS ')
      READ(5,*) NCOHRT
      WRITE(2,103)
  103 FORMAT(1H ,'PLEASE INPUT THE NUMBER OF SEXES')
      READ(5,*) NSEX
      WRITE(2,105)
  105 FORMAT(1H ,'PLEASE INPUT THE NUMBER OF FERTILE AGE COHORTS')
      READ(5,*) NFERT
      WRITE(2,104)
  104 FORMAT(1H ,'THANK YOU')
      CALL MAGIC
      STOP
      END
C-----------------------------------------------------------------------
      SUBROUTINE MAGIC
C-----------------------------------------------------------------------
C     THIS SUBROUTINE SETS THE VALUES OF THE MAGIC VARIABLES
C     WHENEVER THEY ARE NEEDED FOR DIRECT ACCESS READING OR WRITING
C-----------------------------------------------------------------------
      COMMON /A1/ NINREG,NCOHRT,NSEX,NFERT
      COMMON /M/ MAGIC1,MAGIC2,MAGIC3,MAGIC4,MAGIC5,
     *MAGIC6,MAGIC7,MAGIC8,MAGIC9
      MAGIC1 = IFIX((FLOAT(NINREG)/6.)+.875)
      MAGIC2 = NINREG*MAGIC1
      MAGIC3 = 7*MAGIC1 - 2*MAGIC2 + 4
      MAGIC4 = NSEX*MAGIC3
      MAGIC5 = MAGIC4*(NCOHRT + 3)
      MAGIC6 = MAGIC1+1
      MAGIC7 = MAGIC6*(NFERT+1)+(NSEX*MAGIC6)
      MAGIC8 = (NFERT+1)*MAGIC6 +MAGIC1
      MAGIC9 = NCOHRT*NSEX*MAGIC1
      NREC   = 2*MAGIC5+MAGIC7+MAGIC8+MAGIC9
```

A.3 A LISTING OF THE DANREC PROGRAM (CONTINUED)

```
      WRITE(2,100)                                                      DAN00400
  100 FORMAT(//1H ,'THE SPECIFICATION FOR THE DIRECT ACCESS DISK FILE ',DAN00410
     *'(MAGIC PARAMETERS)'/1H ,16('-----')/                             DAN00420
     *1H ,'NUMBER OF RECORDS NEEDED FOR :-')                            DAN00430
      WRITE(2,1) MAGIC1                                                 DAN00440
    1 FORMAT(1H ,6X,'A VECTOR OF FLOWS OR RATES (MAGIC1)',14X,'=',I5)   DAN00450
      WRITE(2,2) MAGIC2                                                 DAN00460
    2 FORMAT(1H ,6X,'A MATRIX OF FLOWS OR RATES (MAGIC2)',14X,'=',I5)   DAN00470
      WRITE(2,3) MAGIC3                                                 DAN00480
    3 FORMAT(1H ,6X,'ONE SEX-AGE COHORT (MAGIC3)',22X,'=',I5)           DAN00490
      WRITE(2,4) MAGIC4                                                 DAN00500
    4 FORMAT(1H ,6X,'ONE AGE COHORT (MAGIC4)',26X,'=',I5)               DAN00510
      WRITE(2,5) MAGIC5                                                 DAN00520
    5 FORMAT(1H ,6X,'ALL AGE COHORTS (MAGIC5)',25X,'=',I5)              DAN00530
      WRITE(2,6) MAGIC6                                                 DAN00540
    6 FORMAT(1H ,6X,'ONE AGE COHORT IN THE BIRTHS SECTION (MAGIC6)',4X, DAN00550
```

FILE: DANREC FORTRAN A LEEDS UNIVERSITY VM/BSE

```
     *'=',I5)                                                           DAN00560
      WRITE(2,7) MAGIC7                                                 DAN00570
    7 FORMAT(1H ,6X,'ALL AGE COHORTS IN THE BIRTHS SECTION (MAGIC7)',3X,DAN00580
     *'=',I5)                                                           DAN00590
      WRITE(2,8) MAGIC8                                                 DAN00600
    8 FORMAT(1H ,6X,'ALL AGE COHORTS IN  FERTILITY SECTION (MAGIC8)',   DAN00610
     *3X,'=',I5)                                                        DAN00620
      WRITE(2,9) MAGIC9                                                 DAN00630
    9 FORMAT(1H ,6X,'THE FINAL POPULATION STACK (MAGIC9)',14X,'=',I5)   DAN00640
      WRITE(2,10) NREC                                                  DAN00650
   10 FORMAT(1H ,6X,'THE WHOLE FILE (NREC)',28X,'=',I5)                 DAN00660
      WRITE(2,101)                                                      DAN00670
  101 FORMAT(1H ,16('-----'))                                           DAN00680
      RETURN                                                            DAN00690
      END                                                               DAN00700
```