Working Paper 205/Computer Manual 2

COMPUTER PROGRAMMES FOR AGGREGATE ACCOUNTS-BASED AND ASSOCIATED FORECASTING MODELS OF POPULATION

J.C. Jenkins and P.H. Rees

School of Geography
University of Leeds
Leeds LS2 9JT.

September 1977

Abstract

Previous work on aggregate accounting models (Rees and Wilson, 1977) and programmes (Rees and Wilson, 1974) is made more general, more flexible and more useful for projection purposes. The models involved and the associated programmes accept variable time series of the main migration, birth and death rates; the flows between the regions in the system of interest and the external world can be dealt with in a variety of ways; the infant and non-infant sections of the accounts can be separated; migrations rather than migrants can be used as initial inputs (for short periods); and final population figures for the regions can be used as constraints on the estimation of the accounts. From the accounts the alternative growth rates matrix model can also be formulated and run.

Part I of the paper outlines these developments. Part II details the models; Part III describes the structure of the computer programme and sets out the instructions for preparing inputs to the programme.

Acknowledgements

The authors would like to thank Alan Wilson, David Illingworth and John Stillwell for their work on earlier versions of an aggregate accounts programme; John Stillwell and Dyl Thomas for supplying data used to test preliminary versions of the revised programmes described in this paper, and all those who have made useful suggestions or comments on the paper's contents.

Thanks are also extended to the Centre for Computer Studies of Leeds University for their efficient and helpful services, and indeed for the opportunity to develop these programmes on the 1906A machine at Leeds.

John Jenkins would also like to thank the Social Science Research Council for funding his postgraduate research in the School of Geography.

Contents

List of Tables

1.

List of Figures

0. Introduction

The aim of this paper is to provide, within a single document, an outline of various accounts-based and associated models of population and some details of a set of three computer programmes which have been developed to implement these models.

We assume at the outset that readers have some basic knowledge of accounting methods of population modelling and, therefore, do not seek to spell out explicitly all our terms of reference, notation and so on. For those without such knowledge the most compact source of information is probably Rees and Wilson (1977, chapters 2-7). Having said this we have attempted to put the paper together in such a way that readers of varying levels of experience of working with accounting techniques and models can refer to sections of particular relevance.

The paper is divided into three parts. As one of the aims of the paper is to update and expand an earlier computer programme user' manual (Rees and Wilson, 1974), the first part is by way of summary of the main modifications made to the original programme. This part of the paper will probably be of particular interest to those who have already used the old programme. The choice of models in each of the three new programmes is then briefly described.

The second part of the paper is intended primarily as a reference section in which the models that can be run in the three programmes are set out in terms of their input requirements, their structure and model equations. The intention in including this section in the paper is to forestall the need of programme users to collect together various other source texts on the models should they wish to check on any particular modelling points. Having said this the paper cannot be used in total isolation - given the varying degrees of familiarity with accounting work it is inevitable that some queries will arise which can only be answered satisfactorily elsewhere. Wherever we think such points are likely to occur we have attempted to direct readers to the most appropriate source of further information.

We should point out, however, that the second part of the paper, and indeed the programmes themselves, are essentially complementary to another paper by the present authors (Rees and Jenkins, 1977). This second paper discusses at some length the more theoretical questions of modelling population at an aggregate level, and the various approaches to the task that can be adopted using an accounting framework.

The third part of the paper, in conjunction with the Appendices, provides most of the details about the programmes themselves - their structure, the input requirements; the output format; and some examples. This part may be consulted as a reference section by readers familiar with the models available in the programmes. It is obviously impractical to reproduce the texts of the three programmes in full in the paper itself, but the authors will be pleased to supply numbered or unnumbered line printer listings of any of the programmes to readers interested in using, adapting or further developing them. Machine readable copies of any of the programmes (probably in the form of paper tape) will also be supplied if required. These, and any listings of interest, will be provided at reproduction and mailing cost. Readers who use the University of Leeds 1906A computer, or have access to it, will be able to produce their own listings of the programmes, and also run their own jobs using the copies of the programmes stored in the Geography programme library. Details of the commands required for these operations are given in Appendix A.

We have sought, therefore, in writing this paper, to produce a document to be read in its own right and a reference manual for users of the computer programmes that are described. We hope that the programmes, and the paper, will prove useful in both a research and a teaching context.

PART I: Some Preliminary Comments

1. Developments of the 1974 Programme

This paper has been produced in response to the need for a reformulation of an earlier computer programme users' manual (Rees and Wilson, 1974). The programme described in that paper related to the aggregate accounts-based model (a.b.m.), and enabled it to be run for a historical or forecasting example. Several limitations of that programme have now become evident and this seems an opportune time to update and extend it by incorporating subsequent developments.

This paper goes beyond a simple reworking of the earlier programme and offers a completely rewritten set of three programmes. Each programme is now structured such that the main body of computing is performed in a series of subroutines rather than in one block of text, as in the old version. It is hoped that this particular arrangement will allow greater flexibility in the use of the programmes by encouraging future users to make further additions and modifications as and when they are required. The aim in producing three programmes has been to offer a wider range of options in terms of input characteristics and models, and, by doing so, to appeal to a wider audience.

Several developments have been introduced into some or all of the three programmes and these will now bw outlined.

1.1 "Auto-Projection".

The SDAT programme of Rees and Wilson (1974) was designed to run a historical or a forecasting version of the aggregate accounts-based model. Each run of the programme related to a single modelling period, and a new block of control and data cards had to be input for each new period. A change of emphasis underlies the development of the set of revised programmes. The historical version of the accounts-based model is still available, but forecasting work is now seen as the main function of the programmes.

Running the historical model is considered to be an integral part of preparing a series of forecasts in order to obtain some appreciation of the trends in various rates, and the programmes have been designed to reflect this. The "auto-projection" option involves the running of a historical base-run of the accounts-based model, followed by a pre-specified number of forecasting periods using either the "Rogers G-matrix" or one of the forecasting versions of the accounts-based model. This whole sequence of historical and forecasting runs is treated as a single operation, therefore, requiring the input of only two sets of control cards, irrespective of the number of forecasting periods being used, the first relating to the historical run and the second to the forecasting series. The data requirements for the historical period are unchanged. Rather than specify a new set of data inputs for each forecasting period, only that data which is to differ from the information

The auto-projection option has been specifically designed in this form to make two particular types of exercise commonly carried out by demographers and population geographers easy to implement. The first is the projection of a set of observed historical trends into the future. This task of applying constant rates of change can now be accomplished without the need to input any data for the forecasting periods since all the information required is produced in the historical run and stored for use in forecasting. The second exercise is the examination of the effects of given changes on future population development. All the components of population change (birth, death and in- and out-migration) can be controlled independently of each other. Consequently the only inputs that need to be specified in such a case are the new values of any particular component that is to be changed. These changes can be introduced at any point in the forecasting series and indeed at as many points as required.

It is still possible to run the accounts-based model for a single historical or forecasting period as in the old programme, but it is thought that such an alternative will now be less attractive given the minimal extra requirements that an auto-projection run may make. A series of single historical runs will, however, probably be a useful precursor to any auto-projection exercise.

The implications of the auto-projection option will probably become clearer when we describe the relationships between the historical and forecasting versions of the accounts-based models in part II of the paper.

1.2 <u>A Flexible Definition of the Regional System.</u>

All population flows that relate to the regions of interest must be considered explicitly if the coverage of any model is to be complete. A region (or set of regions) must not be treated as a closed system if there is migration into or out of that region to and from the rest of the world as this ignores the effect of external migration on the region's population development. Forecasts based on rates defined from an incomplete analysis are not likely to be very trustworthy, and what appear to be small errors in the short term are bound to be amplified as the forecasting period is extended.

It seems virtually impossible to define any truly closed system other than the whole world, even though many population modellers have done otherwise. This does not imply, however, that, where the population system of interest is composed of say the regions of a single country (as is commonly the case), it is always necessary to model the whole world's population development in its entirety to ensure accuracy. Clearly if that were so a significant proportion of the analysis would be irrelevant to modelling the country's population. In addition such a prospect would be daunting both in terms of the effort involved and the data that would be required.

The most reasonable solution seems to be to work with a "semi-closed" system. Given that most population modelling is concerned essentially with only a limited number of regions, the division of the global closed system into two sets of regions - basically those in which we are interested (which we call regional set I) and those in which we are not (which we call regional set O) - would seem realistic. All population flows internal to the main system, as we shall term regional set I, are of interest, as are the flows between the main system and the external regions, as we shall term regional set O. All flows totally within the external system of regions, however, do not need to be considered. Such an arrangement will be termed a "semi-closed" system.

Translating this into less abstract-terms, the only population variables of the external regions that need to be modelled are the migrations to and from the regions ofthis main system. This level of births, deaths and population stocks in the external regions is largely irrelevant, as are the migrations between the regions of the external system, although in certain cases information on these terms may be required for estimation purposes.

In the old programme both the closed and semi-closed regional systems could be used. In the semi-closed case only one external region was allowed, however, and this was, by definition, "the rest of the world". This is a very coarse definition which obscures particular external flows by lumping them all together into one amorphous total. A more refined system is considered preferable in which users, if they so desire, may disaggregate this single rest-of-the-world region into any number of external regions, thereby ensuring that individual external flows of interest can also be dealt with explicitly. Having said this it must be remembered that in total the external regions should still be defined to include the whole of the world outside the main regional system.

The emphasis in the new programmes is on the use of semi-closed regional systems, although various closed-system options are also provided. The data problems implied by using a global closed system lead us to recommend the adoption of a semi-closed system for most modelling exercises. This will avoid, among other things, the task of defining a global population total, which is not only likely to be very inaccurate, but also, because of the nature of the model in the closed form, likely to transfer such inaccuracy into various elements in the final accounts matrix. Of course a non-global closed system may be used in the programmes if users feel that they can realistically define such a thing, but the authors must reiterate their distrust of such regional definitions.

1.3 Semi-Aggregate Models

The aggregate accounts-based model (described in Rees and Wilson, 1977, Chapter 5) uses a set of single average death rates considered applicable to the population of each region. There are, however, two sub-populations being implicitly defined in the aggregate model - the population alive at the start of the modelling period, and that born in the course of the period of interest. Throughout this paper we shall refer to the first of these populations as the "non-infant" population, and to the second as the "infant". This is simply by way of a shorthand convention and we recognize that many persons defined as non-infants here would be considered infants in the normal sense of the word. This dichotomy is essentially a crude age division of the population into those over and under x years of age at the end of the modelling period, where x is the length of the period in years.

If the modelling period is one year, for example, then the use of an average death rate will tend to underestimate the deaths of those born in that year since this group is subject to a higher death rate than the average on account of the relatively large numbers of infants dying within a few days of their birth. On the other hand, if the modelling period is say five years, the use of an average death rate will tend to overestimate the deaths of under fives as this group is subject to a lower death rate than the average. Consequently it would seem useful to make this implicit division of the population into two sub-populations fully explicit in the model by compiling demographic accounts for the two groups on a completely independent basis. This would produce a model intermediate between the fully aggregate and the age-disaggregated levels of interest. This "semi-aggregate" model recognises that, for one sub-section of the population, age is an important determinant of death rates.

In such a model a separate infant death rate will be calculated and applied to those born in the period. The only change required in the data inputs is that, in the historical case, the deaths total be replaced by two sub-totals (one for each sub-group of the population), and that, in the forecasting case, instead of a single death rate two are now needed. As the deaths information is not generally available in an appropriate form, a short programme has been written which produces estimates on the basis of known deaths data. Details of this programme may be found in Appendix B. The semi-aggregate model is spelled out in greater detail in section 4 of the paper.

1.4 M-Data Accounts

All the published examples to date of the application of demographic accounting models, derived from the well-defined set of principles most fully summarised in Rees and Wilson (1977), have been for various British regional systems. The methodology that has been developed in those examples reflects the dominant influence of the British experience. The inputs required by the model have been specified in terms of the information that is available in the British case, and consequently the earlier version of the computer programme was devised under the same terms of reference. Existing accounting models require data on migration, for example, in terms of the number of people whose location at the start of the period was different from that at the end of the period. This is exactly the information provided by the retrospective migration question (i.e.: "Where did you live x years ago?") asked at the most recent British censuses. Since it is now the convention to use the symbol K to represent the various population terms referred to in accounting models, this type of migration information, which may be used directly in the accounts matrix, will be referred to as being of the "K-data" form.

A recent review by one of the present authors, who has been considering the possibilities of modelling population at an international scale in Western Europe (Jenkins 1976), suggests that only a few nations other than Britain ask the type of retrospective census migration question referred to earlier. A significant number prefer to collect statistics on migration through a continuous registration system. This sort of system does not provide K-data, but rather what will be referred to as "M-data".

The crucial difference between the two types of data lies in what is actually being recorded in each case. K-data refers to persons, who are migrants in an accounting sense (that is their initial location differs from their final location). On the other hand M-data refers to the event of moving and as such is a count of migrations. This distinction between migrants and migrations is an important one in modelling population, as Courgeau (1973) has pointed out.

In many respects the continuous counting of moves provides a more complete and accurate coverage of the phenomenon than data from census questions. The shortcomings, and their implications in terms of providing data for modelling, of both types of system of recording migration have been discussed by the authors in some detail elsewhere (see Rees 1977a, Jenkins 1976). Briefly M-data is the sum of the appropriate migration terms from each of the four sub-matrices of the full accounts matrix, plus a "migration surplus" term:-

$$M^{*(i)*(j)} = K^{\varepsilon(i)\sigma(j)} + K^{\varepsilon(i)\delta(j)} + K^{\beta(i)\sigma(j)} + K^{\beta(i)\delta(j)} + M^{*(i)*(j)\mathrm{SURPLUS}} \qquad (1.1)$$

where

$M^{*(i)*(j)}$ = the sum of all the migrations from region i to region j in a period;

= $\sum_{\alpha\omega}\sum M^{\alpha(i)\omega(j)}$ where α refers to lifestate prior to migration and ω to lifestate after migration. α can take on two values, either β, representing birth, or ε, representing existence; ω can take on two values, σ, representing survival, or δ, representing death;

$K^{\varepsilon(i)\sigma(j)}$ = the number of persons in existence in region i at the beginning of the period who survive in region j at the end;

$K^{\varepsilon(i)\delta(j)}$ = the number of persons in existence in region i at the start of the period who die in region j before the end of the period;

$K^{\beta(i)\sigma(j)}$ = the number of persons born in region i who survive in region j at the end of the period;

$K^{\beta(i)\delta(j)}$ = the number of persons born in region i who die in region j before the end of the period;

$M^{*(i)*(j)}$ SURPLUS = the migrations between region i and region j irrespective of initial lifestate or final lifestate, surplus to locating persons in the previous four variables.

The superscript * indicates summation over the superscript normally located in that position.

Equation (1.1) shows that the M-data term includes both of the migrant terms ($K^{\varepsilon(i)\sigma(j)}$ and $K^{\beta(i)\sigma(j)}$) required as inputs in the existing versions of the accounts-based models, but also counts the migrants who die after moving. The migration surplus term is effectively a measure of the difference between the number of migrants and the number of migrations. This difference results from persons making more than one migration in a given period. The relationship between K- and M-data has been explored in Rees (1974) in some detail.

Equation (1.1) also shows that M-data can be expressed in terms of its K-data equivalents. As a result, a set of manipulations could be designed so that estimates of the $K^{\varepsilon(i)\sigma(j)}$ and $K^{\beta(i)\sigma(j)}$ terms required as inputs to the existing accounts-based model could be arrived at from the $M^{*(i)*(j)}$ figures in examples where the appropriate K-data is not available. Such an approach, while allowing existing models and programmes to be used in their present form with only small changes, would be unproductive in that the task of the model would be to virtually recreate the original data. Instead the models were reformulated, using the accounting principles defined in Rees and Wilson (1977), to accommodate this alternative data source commonly found outside Britain.

The resulting M-data model has been specified in the semi-aggregate form, and has other features which distinguish it from the other existing models. The main one of these is that as a result of the nature of the available migration date a final population total will be implicitly known in the historical case from the equation:-

$$K^{*(*)\sigma(i)} = K^{\varepsilon(i)*(*)} + K^{\beta(i)*(*)} - K^{*(*)\delta(i)} + \sum_{j\neq i} M^{*(j)*(i)} - \sum_{j\neq i} M^{*(i)*(j)} \qquad (1.2)$$

where

$K^{*(*)\sigma(i)}$ = the end of period number of persons surviving in region i or the final population of region i;

$K^{\varepsilon(i)*(*)}$ = the number of persons in existence in region i at the start of the period or the initial population;

$K^{\beta(i)*(*)}$ = the number of persons born in region i during the period irrespective of final lifestate or location or the total births in region i;

$K^{*(*)\delta(i)}$ = the number of persons who have died in region i in the period or the total of deaths in region i.

This final total acts, together with the migration totals, as an extra constraint on the values that the elements of the accounts matrix may take.

One outstanding problem should be noted at this point in relation to the migration surplus term in equation (1.1). A numerical estimate of this term is required so that the M-data term can be converted into its K-data components. The nature of the problem is, however, rather more complex than would appear from what we have already said. Given that a satisfactory solution to it has not yet been found, although some progress has been made in Rees (1977a), we do not wish to confuse readers with unnecessary detail and so, in the absence of any other information, the assumption will be made at present that there is no multiple migration, and therefore that the migration surplus is zero. For a short modelling period, say of one year, such an assumption is probably fairly reasonable, but the longer the period becomes the more unreasonable is the assumption. The M-data model is described in more detail in section 5 of the paper.

A forecasting version of the M-data model has not been designed. The existing forecasting models are suitable as they stand, given that it is up to the user to provide the migration data in the form he wishes, rather than having to adopt the information in the particular form in which it is available. Any type of historical accounts-based model, whether it be K- or M-data based, aggregate or semi-aggregate, will produce a standard accounts matrix. This matrix, and the rates matrices defined from it, are the bases for specifying inputs to the forecasting model, hence the lack of any distinction regarding data types in the forecasting case.

1.5 Observed population constraints

The row totals of the existence (ε) section of the accounts matrix give the initial populations of each region, and those of the births (β) section the total births in each region. Similarly the column totals of the survival (σ) section of the accounts matrix give the end-of-period populations of each region, and those of the deaths (δ) section the total deaths. In the M-data version of the historical model,as we noted in section 1.4, all these four sets of totals, including the end-of-period populations, are used in the model equations, either explicitly or implicitly, to ensure that the elements of the accounts matrix sum both row- and column-wise to the observed figures.

In the K-data version, however, because of the nature of the migration data inputs being used, equation (1.2) cannot be completed as, returning to equation (1.1), only the migration and survival terms are known rather than total migration. The implication of this is that, in this form of the model, end-of-period population totals are not used as constraints on the values of elements in the estimated accounts matrix. Summing the columns of the survival half of the accounts matrix, therefore, produces model estimates of the final populations. These can be usefully compared with the observed populations as a means of assessing the accuracy of the model's results.

An optional routine has been added which calculates various statistics to compare the two sets of final population totals. This routine then uses the observed totals in place of the estimates to adjust the elements of the accounts matrix accordingly, through a method commonly used in urban and regional analysis, that of balancing factors. A useful summary of balancing factor methods is to be found in MacGill (1975), and Illingworth (1976) gives examples of their use in a demographic context.

In general terms balancing factors are, to quote MacGill's working definition, "multiplicative terms...that can be used to adjust a set of initial values for the components of some system of interest to produce a set of estimates for these components, based firmly on the initial values, which agree with whatever pre-specified aggregate totals...we intend they should satisfy." Translating this into less general terms more relevant to our particular problem, balancing factors may be used to adjust elements of the accounts matrix estimated by the model so that the column totals of the survival half of the matrix are equal to the observed end-of-period populations. In doing this they also ensure that the remaining row and column totals of the matrix still equal the appropriate observed totals that the model itself used as constraining information (that is, the initial populations, and births and deaths totals). We shall consider the closed system case first to illustrate the nature of the balancing factor operations, and then discuss the additional difficulties raised in the semi-closed system example.

We employ a doubly-constrained balancing factor routine in the programmes in which both row and column constraints of the accounts matrix must be satisfied. This doubly-constrained adjustment, used to produce a revised* accounts matrix ($\hat{K}^{\alpha(i)\omega(j)}$), can be expressed as:-

$$\hat{K}^{\alpha(i)\omega(j)} = A_i^{\alpha} . B_j^{\omega} . K^{\alpha(i)\omega(j)} \tag{1.3}$$

where α and ω are general labels for initial (ε or β) and final (σ or δ) states respectively; and A_i^{α} and B_j^{ω}, the balancing factors, are given by:-

$$A_i^{\alpha} = \frac{\hat{K}^{\alpha(i)*(*)}}{\sum\limits_{\omega j}\sum B_j^{\omega} . K^{\alpha(i)\omega(j)}} \tag{1.4}$$

and

$$B_j^{\omega} \quad \frac{\hat{K}^{*(*)\sigma(i)}}{\sum\limits_{\alpha i}\sum A_i^{\alpha} . K^{\alpha(i)\omega(j)}} \tag{1.5}$$

These equations for the balancing factors are arrived at by considering their functions. The A_i^{α} terms ensure that the row sums are satisfied:-

$$\sum_j \hat{K}^{\varepsilon(i)\sigma(j)} + \sum_j \hat{K}^{\varepsilon(i)\delta(j)} = \hat{K}^{\varepsilon(i)*(*)} \tag{1.6}$$

$$\sum_j \hat{K}^{\beta(i)\sigma(j)} + \sum_j \hat{K}^{\beta(i)\delta(j)} = \hat{K}^{\beta(i)*(*)} \tag{1.7}$$

or more generally:-

$$\sum_{\omega j}\sum \hat{K}^{\alpha(i)\omega(j)} = \hat{K}^{\alpha(i)*(*)} \tag{1.8}$$

---

* Throughout this section we shall adopt the following notational conventions:-

a. for the accounts matrix, and its row and column totals, as estimated by the model, we shall use the ordinary K label;

b. for the revised matrix, and its row and column totals (and consequently the observed row and column constraints) we shall use a $\hat{K}$ label.

Hence equation (1.4) is arrived at by substituting the right hand side of equation (1.3) into equation (1.8) and rearranging the terms. The B_j^{ω} terms, which ensure that parallel equations for the column sums are satisfied, are arrived at in a similar way by substituting the right hand side of equation (1.3) into equation (1.9), which is:-

$$\hat{K}^{*(*)\omega(j)} = \sum\sum_{\alpha i} \hat{K}^{\alpha(i)\omega(j)} \tag{1.9}$$

and rearranging the terms. An important condition in this form of routine is that the sum of all the row constraints must equal the sum of all the column constraints. This may be expressed as:-

$$\sum_{i} \hat{K}^{\varepsilon(i)*(*)} + \sum_{i} \hat{K}^{\beta(i)*(*)} = \sum_{j} \hat{K}^{*(*)\sigma(j)} + \sum_{j} \hat{K}^{*(*)\delta(j)} \tag{1.10}$$

or more generally:-

$$\sum\sum_{\alpha i} \hat{K}^{\alpha(i)*(*)} = \sum\sum_{\omega j} \hat{K}^{*(*)\omega(j)} \tag{1.11}$$

So, in using the constraint routine in the suite of programmes, the user should ensure that the total of the row constraints is equal to the total of the column constraints. Otherwise the set of balancing factor equations will not converge. Particular care must be given to this in the semi-closed case.

In addition to the specific developments that have already been outlined the number of options available in the forecasting models in terms of input requirements has been extended, and these extensions are discussed in Part II of the paper in which all the models are described.

2. The Revised Programmes

Rather than incorporate all the developments described in section 1 into one large reformulated programme we decided to produce a series of separate programmes. This decision was a function of the fact that some of the developments were intended as alternatives rather than additions to the "old" model and programme. Rather than choose options within a single programme, therefore, the choice will initially be between three programmes as we felt that few users would require all the versions of the accounts-based model on offer, at least in a single exercise. This approach has avoided the necessity of writing and testing a long and relatively complex programme. The benefits of shorter, simpler programmes will also accrue to users in terms of lower demands on their computing budgets, easier prospects for changing programmes to suit their own needs and so on.

Three of the developments that have been discussed can be placed in the category of additions, and as such have been incorporated into all three programmmes. These are the auto-projection mode of analysis, the flexible regional system and the end-of-period constraint option. As well as the new versions of the accounts-based model, a forecasting model which may be derived from the accounts (see Rees, 1976, for an example) has been incorporated as an alternative in the three programmes. This latter model, which uses a matrix of growth rates, will be called here the "Rogers G-matrix forecasting model" after Rogers (see Rogers, 1966) who first developed it. This model is described in more detail in section 6 of the paper.

The three programmes that have been produced are named MULTIPOP1, MULTIPOP2 and MULTIPOP3 (MULTI-regional POPulation accounting and forecasting programme, version 1, version 2, version 3) The MULTIPOP1 programme offers the choice of three historical accounts-based models (the original aggregate model, and the semi-aggregate K-data and M-data models) with one forecasting model (the Rogers G-matrix). The MULTIPOP2 programme offers the fully aggregate accounts-based model in both its historical and forecasting forms, together with the Rogers G-matrix as a forecasting alternative. The MULTIPOP3 programme offers the two semi-aggregate models (K- and M-data) in historical form, and the semi-aggregate forecasting version of the accounts-based model, again with the Rogers G-matrix model as a forecasting alternative. The model options that are available in each of the programmes are summarised for easy reference in Table 1.

The range of models offered in each programme has been decided upon on the basis that transfer from an aggregate historical to a semi-aggregate forecasting model, or from a semi-aggregate historical to an aggregate forecasting model is rather illogical. The first transfer would involve making assumptions about infant deaths rate in the forecasting period that information from the historical model does not provide. The second would involve discarding information on infant death rates that had been available in historical runs of the model.

Table 1: Features contained in the three alternative programmes.

HISTORICAL CASES	MULTIPOP1	MULTIPOP2	MULTIPOP3
A. Aggregate Model			
K-data a.b.m. (section 3.1)	YES	YES	NO
B. Semi-Aggregate Model			
K-data a.b.m. (section 4)	YES	NO	YES
M-data a.b.m. (section 5)	YES	NO	YES
FORECASTING CASES			
A. Aggregate a.b.m. (section 3.2)	NO	YES	NO
B. Semi-Aggregate a.b.m. (section 4)	NO	NO	YES
C. Rogers G-Matrix Model* (section 6)	YES	YES	YES

* This model may be used in either an aggregate or a semi aggregate context.

PART II: DESCRIPTION OF THE MODELS

In this part of the paper the aim is to outline the basic structure and equations of the models referred to in Part I, and made available in the computer programmes. The aggregate accounts-based model of the original programme is described in the greatest detail as the majority of users will be most familiar with it. Many of the points raised in relation to that model, however, such as the nature of the relationship between historical and forecasting versions, are applicable to the other accounts-based models in later sections, and so are not necessarily repeated.

3. The Aggregate Accounts-Based Model

3.1 Historical Mode

The version of this model used in the programmes is that described in Rees and Wilson (1977, chapter 5) with minor changes. The structure of the model, illustrated in Figure A, is as follows:-

a. KNOWN INFORMATION. The following population data is available from existing sources, and may, therefore, be used directly as inputs to the model:-

$K^{\varepsilon(i)*(*)}$	Initial population
$K^{\beta(i)*(*)}$	Total births
$K^{*(2)\delta(i)}$	Total deaths
$K^{\varepsilon(i)\ (j)}$	Surviving non-infact migrants

b. SURVIVING INFANT MIGRANTS. This term is not directly available and must be estimated, either exogenously and thereby treated as an input, or inside the modelling routine of the programme. In either case the same equation will probably be used. This assumes that infants migrate at the same rate as non-infants:-

$$K^{\beta(i)\sigma(j)} = 0.5\ \frac{K^{\varepsilon(i)\sigma(j)}}{K^{\varepsilon(i)*(*)}}\ K^{\beta(i)*(*)} \qquad (3.1)$$

c. POPULATION AT-RISK OF GIVING BIRTH (INITIAL ESTIMATE). Although this term is calculated in the historical case, it is actually redundant in the model. Since all the population terms needed to estimate it (see equation 3.15 later) are not known at this stage, an initial estimate of its value must be made using those which are:-

$$\hat{K}^{Bi*} = K^{\varepsilon(i)*(*)} + 0.5 \sum_{j\neq i} K^{\varepsilon(j)\sigma(i)} - 0.5 \sum_{j\neq i} K^{\varepsilon(i)\sigma(j)} \qquad (3.2)$$

Note that members of the infant population are not considered to be at-risk of giving birth.

Figure A: The Structure of the Historical Aggregate A.B.M.

ASSEMBLE KNOWN DATA

${}_K^{\varepsilon(i)*(*)}$

${}_K^{\varepsilon(i)\sigma(j)}$

${}_K^{\beta(i)*(*)}$

${}_K^{*(*)\delta(i)}$

CALCULATE SURVIVING INFANT MIGRANTS

${}_K^{\beta(i)\sigma(j)}$ (3.1)

INITIAL ESTIMATES OF AT-RISK POPULATIONS

${}_K^{\hat{B}i*}$ (3.2)

${}_K^{\hat{D}*i}$ (3.4)

CALCULATE BIRTH AND DEATH RATES

b^{i*} (3.3)

d^{*i} (3.5)

RECALCULATE AT-RISK POPULATIONS

${}_K^{\hat{D}*i}$ (3.14)

${}_K^{\hat{B}i*}$ (3.15)

CALCULATE MIGRANT DEATHS

${}_K^{\varepsilon(i)\delta(j)}$ $j\varepsilon I$ (3.6)

${}_K^{\beta(i)\delta(j)}$ $j\varepsilon I$ (3.7)

${}_K^{\varepsilon(i)\delta(j)}$ $j\varepsilon O$ (3.8)

${}_K^{\beta(i)\delta(j)}$ $j\varepsilon O$ (3.9)

CALCULATE NON-MIGRANT TERMS

${}_K^{\beta(i)\delta(i)}$ (3.10)

${}_K^{\varepsilon(i)\delta(i)}$ (3.11)

${}_K^{\varepsilon(i)\sigma(i)}$ (3.12)

${}_K^{\beta(i)\sigma(i)}$ (3.13)

CONVERGENCE TESTS → NOT CONVERGED

CONVERGED

CALCULATE FINAL POPULATIONS

${}_K^{*(*)\sigma(i)}$

d. BIRTH RATE. This rate, derived from the at-risk population, is redundant in the historical version of the model like the population itself. It does seem valuable, however, to estimate it as a prerequisite to the forecasting analysis where it is used in the model equations:-

$$b^{i*} = \frac{K^{\beta(i)*(*)}}{{}_{K}\hat{B}^{i*}} \tag{3.3}$$

e. POPULATION AT-RISK OF DYING (INITIAL ESTIMATE). Unlike the at risk population for births this term is required in the historical case. An initial estimate is given by:-

$$\begin{aligned} {}_{K}\hat{D}^{*i} = {} & K^{\varepsilon(i)*(*)} + 0.5 \sum_{j\neq i} K^{\varepsilon(j)\sigma(i)} \\ & - 0.5 \sum_{j\neq i} K^{\varepsilon(i)\sigma(j)} + 0.5\, K^{\beta(i)*(*)} \\ & + 0.25 \sum_{j\neq i} K^{\beta(j)\sigma(i)} - 0.25 \underset{j\neq i}{} K^{\beta(i)\sigma(j)} \end{aligned} \tag{3.4}$$

f. DEATH RATE. Again required in the historical case:-

$$d^{*i} = \frac{K^{*(*)\delta(i)}}{{}_{K}\hat{D}^{*i}} \tag{3.5}$$

g. MIGRANT DEATHS. These terms are calculated as minor flows. The number of migrants who die in the main system region is estimated separately from that of migrants dying in the external regions. In the former case the death rate of the destination region is used:-

$$K^{\varepsilon(j)\delta(i)} = 0.5\, \frac{d^{*i}\; K^{\varepsilon(j)\sigma(i)}}{1-0.25\, d^{*i}} \quad i\varepsilon I;\ j\varepsilon I,0 \tag{3.6}$$

$$K^{\beta(j)\delta(i)} = 0.25\, \frac{d^{*i}\; K^{\beta(j)\sigma(i)}}{1-0.125\, d^{*i}} \quad i\varepsilon I;\ j\varepsilon I,0 \tag{3.7}$$

where all j reasons (both main and external) are to be considered. In the latter case the death rate of the origin region is used:-

$$K^{\varepsilon(j)\delta(i)} = 0.5 \frac{d^{*j} K^{\varepsilon(j)\sigma(i)}}{1-0.25d^{*j}} \quad i\varepsilon O;\ j\varepsilon I \tag{3.8}$$

$$K^{\beta(j)\delta(i)} = 0.25 \frac{d^{*j} K^{\beta(j)\sigma(i)}}{1-0.125d^{*j}} \quad i\varepsilon O;\ j\varepsilon I \tag{3.9}$$

This distinction between the two sets of regions is made because of the difficulty in defining reasonable death rates for the external regions. In order to define such rates it would be necessary to model the populations of the external regions explicitly, a task we have sought to avoid for reasons explained earlier.

h. NON-MIGRANTS. The first of the non-migrant terms is calculated as a minor flow, and the remaining three as major flows using accounting equations:-

$$K^{\beta(i)\delta(i)} = 0.5\ d^{*i} \frac{(K^{\beta(i)*(*)} - 0.5 \sum_{j\neq i} K^{\beta(i)\sigma(j)} - 0.75 \sum_{j\neq i} K^{\beta(i)\delta(j)})}{1 + 0.25\ d^{*i}} \tag{3.10}$$

$$K^{\varepsilon(i)\delta(i)} = K^{*(*)\delta(i)} - \sum_{j\neq i} K^{\varepsilon(j)\delta(i)} - \sum_{j\neq i} K^{\beta(j)\delta(i)} - K^{\beta(i)\delta(i)} \tag{3.11}$$

$$K^{\varepsilon(i)\sigma(i)} = K^{\varepsilon(i)*(*)} - K^{\varepsilon(i)\delta(i)} - \sum_{j\neq i} K^{\varepsilon(i)\delta(j)} - \sum_{j\neq i} K^{\varepsilon(i)\sigma(j)} \tag{3.12}$$

$$K^{\beta(i)\sigma(i)} = K^{\beta(i)*(*)} - K^{\beta(i)\delta(i)} - \sum_{j\neq i} K^{\beta(i)\sigma(j)} - \sum_{j\neq i} K^{\beta(i)\delta(j)} \tag{3.13}$$

All the elements of the accounts matrix have now been calculated using a set of initial estimates for the at-risk populations.

The next stage of the modelling procedure is the convergence testing routine. These tests involve monitoring the variation of particular accounts elements from one iteration of the model to the next, and converting this variation into a relative index. If the value of all such indices is less than a pre-specified value then the model is said to have converged, that is, it has reached a point beyond which further iterations will not significantly change the results obtained.

In all the models in this paper all the deaths elements of the accounts matrix are tested for convergence, including the $K^{\varepsilon(i)\delta(i)}$ terms, which, as major flows, were previously not tested in the original programme. It was observed, however, that in that programme, when the model had been said to have converged on the basis of testing only the minor flows, the major flows still continued to vary. This is a function of their being significantly larger than the minor flows in most cases. Since major flows are now being tested, the convergence procedure is carried out after the accounting equations rathdr than before.

i. FULL AT-RISK POPULATIONS. If any of the accounts elements do vary significantly, as they are bound to after only one iteration (given that all the items in the previous iteration will have been zero), the at risk populations are recalculated using the full equations:-

$$\begin{aligned}\hat{K}^{D*i} = {} & K^{\varepsilon(i)*(*)} + 0.5 \sum_{j\neq i} K^{\varepsilon(j)\sigma(i)} - 0.5 \sum_{j\neq i} K^{\varepsilon(i)\sigma(j)} \\ & - 0.5\, K^{\varepsilon(i)\delta(i)} + 0.25 \sum_{j\neq i} K^{\varepsilon(j)\delta(i)} - 0.75 \sum_{j\neq i} K^{\varepsilon(i)\delta(j)} \\ & - 0.25\, K^{\beta(i)\delta(i)} + 0.125 \sum_{j\neq i} K^{\beta(j)\delta(i)} - 0.375 \sum_{j\neq i} K^{\beta(i)\delta(j)} \\ & + 0.5\, K^{\beta(i)*(*)} - 0.25 \sum_{j\neq i} K^{\beta(i)\sigma(j)} + 0.25 \sum_{j\neq i} K^{\beta(j)\sigma(i)} \qquad (3.14)\end{aligned}$$

$$\hat{K}^{Bi*} = K^{\varepsilon(i)*(*)} + 0.5 \sum_{j\neq i} K^{\varepsilon(j)\sigma(i)} - 0.5 \sum_{j\neq i} K^{\varepsilon(i)\sigma(j)}$$

$$-0.5\, K^{\varepsilon(i)\delta(i)} + 0.25 \sum_{j\neq i} K^{\varepsilon(j)\delta(i)}$$

$$- 0.75 \sum_{j\neq i} K^{\varepsilon(i)\delta(j)} \qquad (3.15)$$

The form of full at-risk equations may vary between different versions of the aggregate model as substitutions can be made which allow the two at risk populations to be estimated from combinations of other terms.

The recalculated at-risk populations are then used in place of the initial or previous estimates in equations (3.3) and (3.5). Equations (3.6) to (3.13) are then repeated, testing for convergence after each iteration of the model. This cycle is performed as many times as necessary until convergence is finally achieved.

3.2 <u>Forecasting Mode</u>

The structural similarity of the historical and forecasting versions of the accounts-based model is shown by comparing Figures A and B. Indeed the heart of both models is identical. This central section is made up of the minor flow and accounting equations (3.6 to 3.13) in both cases. The implication of this is that the information required in these equations to estimate unknown terms in the accounts will, therefore, also be identical. The most basic difference between the two versions lies in the nature of the manipulations performed on the available inputs to produce the basic information required in this set of central model equations. Figure C is an attempt to clarify this statement and to demonstrate the relationship between the two modes of the model.

FIGURE B: THE STRUCTURE OF THE AGGREGATE FORECASTING A.B.M.

ASSEMBLE KNOWN DATA
$_K\varepsilon(i)*(*)$

ASSEMBLE ESTIMATED BIRTH AND DEATH RATES
b^{i*}; d^{*i}

ASSEMBLE ESTIMATED INPUTS FOR MIGRANT CALCULATIONS
(SEE TABLE 2 FOR OPTIONS)

CALCULATE SURVIVING NON-INFANT MIGRANTS
(SEE TABLE 2 FOR EQUATIONS)
$_K\varepsilon(i)\sigma(j)$

INITIAL ESTIMATE OF POPULATIONS AT-RISK OF GIVING BIRTH
$_K\hat{B}i*$ (3.2)

CALCULATE BIRTH TOTALS
$_K\beta(i)*(*)$ (3.20)

RECALCULATE AT-RISK POPULATIONS
$_K\hat{B}i*$ (3.15)

CALCULATE SURVIVING INFANT MIGRANTS
(SEE TABLE 2 FOR EQUATIONS)
$_K\beta(i)\sigma(j)$

ESTIMATE POPULATIONS AT-RISK OF DYING

FIRST ITERATION
$_K\hat{D}*i$ (3.4)

SUBSEQUENT ITERATIONS
$_K\hat{D}*i$ (3.14)

CALCULATE DEATH TOTALS
$_K*(*)\delta(i)$ (3.21)

CALCULATE MIGRANT DEATHS
(as in historical model - 3.6 to 3.9)

CALCULATE NON-MIGRANT TERMS
(as in historical model - 3.10 to 3.13)

CONVERGENCE TESTS

CONVERGED → NOT CONVERGED

CALCULATE FINAL POPULATIONS
$_K*(*)\sigma(i)$

FIGURE C : INPUTS TO THE HISTORICAL AND FORECASTING VERSIONS OF THE AGGREGATE A.B.M.

HISTORICAL

KNOWN TERMS

${}_K\epsilon(i)*(*)$ @

${}_K\beta(i)*(*)$ @

${}_K*(*)\delta(i)$ @

${}_K\epsilon(i)\sigma(j)$ @

ESTIMATED TERMS

${}_K\beta(i)\sigma(j)$ &

b^{i*} £

d^{*i} £

ESTIMATED VARIABLES

${}_K\hat{B}\ i*$ £

${}_K\hat{D}\ *i$ £

FORECASTING

KNOWN TERMS

${}_K\epsilon(i)*(*)$ @

ESTIMATED TERMS

${}_K\epsilon(i)\sigma(j)$ &

${}_K\beta(i)\sigma(j)$ &

${}_K\beta(i)*(*)$ £

${}_K*(*)\delta(i)$ £

b^{i*} @

d^{*i} @

ESTIMATED VARIABLES

${}_K\hat{B}\ i*$ £

${}_K\hat{D}\ *i$ £

ESTIMATED INPUTS @
(see table 2)

TERMS REQUIRED IN THE MODEL EQUATIONS

${}_K\epsilon(i)*(*)$

${}_K\beta(i)*(*)$

${}_K*(*)\delta(i)$

${}_K\epsilon(i)\sigma(j)$

${}_K\beta(i)\sigma(j)$

b^{i*}

d^{*i}

MINOR FLOW AND ACCOUNTING EQUATIONS
(3.6) to (3.13)

@ EXOGENOUS INPUT

& Choice between ENDOGENOUS and EXOGENOUS ESTIMATION

£ ENDOGENOUSLY CALCULATED

In the forecasting case the only known information will be the initial base population totals. Data will not be available for the total number of births and deaths or for the surviving migrants as in the historical case, in addition to those terms that were previously unknown and had to be estimated. In developing programmes to implement the forecasting model the decision must be made as to how a value for each of these unknown is to be calculated. Are they to be estimated exogenously, or insde the programme from known data? How many of the alternative ways of estimating the unknowns needed to be included?

On the second point, we show (in Rees and Jenkins 1977), the very wide variety of alternatives that do exist, and, although we have attempted to include some of those we felt to be the most useful and commonly required in the programmes, it seems inevitable that users may wish to use some of the alternatives not inc luded. We hope that they are able to adjust one of the programmes here to suit their own requirements.

A variety of options for the estimation of the migration terms has been provided. In the first option the surviving migrant flows are to be estimated exogenously and input directly as required in the model equations. In the remaining options the flows are calculated internally from transition rates that are input to the programme and applied to the appropriate at-risk population:-

$$_{K}^{\varepsilon(i)\sigma(j)} = {}_{H}^{\varepsilon(i)\sigma(j)} \; {}_{K}^{\varepsilon(i)*(*)} \tag{3.16}$$

$$_{K}^{\beta(i)\sigma(j)} = {}_{H}^{\beta(i)\sigma(j)} \; {}_{K}^{\beta(i)*(*)} \tag{3.17}$$

These equations apply to the closed system case. The different options that will now be outlined are variations on this theme.

The variations relate essentially to the manner of estimating the surviving migrant flows to and from the external regions in the semi-closed system case.

In the semi-closed system case there will be no initial population or birth totals defined for the external regions so that equations (3.16)and (3.17) cannot be applied as they stand. One option is to apply the equations only when region i is one of the main system regions:-

$$K^{\varepsilon(i)\sigma(j)} = H^{\varepsilon(i)\sigma(j)} K^{\varepsilon(i)*(*)} \quad i\varepsilon I; \quad j\varepsilon I,O \tag{3.16a}$$

$$K^{\beta(i)\sigma(j)} = H^{\beta(i)\sigma(j)} K^{\beta(i)*(*)} \quad i\varepsilon I; \quad j\varepsilon I,O \tag{3.17a}$$

A distribution, rather than a transition, rate ($\overline{H}$ rather than H) is produced in the semi-closed historical model by dividing each accounts element by its row totals for the external i regions. This distribution rate may be used in this option to obtain surviving migrant flows from external regions to main system regions. This rate is applied to the equivalent of the initial population, in this case the migrant total:-

$$K^{\varepsilon(i)\sigma(j)} = \overline{H}^{\varepsilon(i)\sigma(j)} \sum_{j\neq i} K^{\varepsilon(i)*(j)} \quad i\varepsilon O; \quad j\varepsilon I,O \tag{3.18}$$

$$K^{\beta(i)\sigma(j)} = \overline{H}^{\beta(i)\sigma(j)} \sum_{j\neq i} K^{\beta(i)*(j)} \quad i\varepsilon O; \quad j\varepsilon I,O \tag{3.19}$$

The third migration option again uses equations (3.16a) and (3.17a), but allows the surviving migrant terms calculated in the second option by equation (3.18) and (3.19) to be estimated exogenously and read in directly as flows.

The fourth option is parallel to the second except that "emigration" flows (that is from the main system regions to external regions) are estimated exogenously and input as flows instead of being calculated in equations (3.16a) and (3.17a). These two equations are, therefore, more accurately specified in this option as:-

$$K^{\varepsilon(i)\sigma(j)} = H^{\varepsilon(i)\sigma(j)} \; K^{\varepsilon(i)*(*)} \quad i\varepsilon I; \quad j\varepsilon I \tag{3.16b}$$

$$K^{\beta(i)\sigma(j)} = H^{\beta(i)\sigma(j)} \; K^{\beta(i)*(*)} \quad i\varepsilon I; \quad j\varepsilon I \tag{3.17b}$$

and now are used to estimate surviving migrants within the main system of regions only "Immigration" flows (that is from external regions to main system regions) are again calculated using equations (3.18) and (3.19).

The fifth option is parallel to the third, but again the "emigration" flows are to be input directly in that form, as are the "immigration" flows. Equations (3.16b) and (3.17b) are again used to estimate the remaining terms. All the options for producing the surviving migrant terms are summarised in Table 2 for easy reference.

The discussion of ways of estimating the surviving migrant terms bring another difference between the historical and forecasting versions of the accounts-based model to light. We have used the total births term in the equations to produce the surviving infant migrant estimates, but in the forecasting case this is not a known variable.

In the forecasting version of the model births, and deaths, information must be input in the form of rates, and the births and deaths totals are then calculated from rearranged equations (3.3)and (3.5):-

$$K^{\beta(i)*(*)} = b^{i*} \; \hat{K}^{Bi*} \tag{3.20}$$

$$K^{*(*)\delta(i)} = d^{i*} \; \hat{K}^{D*i} \tag{3.21}$$

The estimation of the surviving infants must, therefore, be included as part of the iterative cycle in which the births total it requires is also calculated.

TABLE 2: THE OPTIONS FOR ESTIMATING THE SURVIVING MIGRANT TERMS IN THE FORECASTING MODELS

	EXOGENOUSLY-ESTIMATED INPUTS		INTERNAL EQUATIONS
OPTION 1	$K^{\varepsilon(i)\sigma(j)}$, $K^{\beta(i)\sigma(j)}$	$i\varepsilon I,0$; $j\varepsilon I,0$	
OPTION 2	$H^{\varepsilon(i)\sigma(j)}$, $H^{\beta(i)\sigma(j)}$	$i\varepsilon I$; $j\varepsilon I,0$	3.16a
	$\overline{H}^{\varepsilon(i)\sigma(j)}$, $\overline{H}^{\beta(i)\sigma(j)}$, $\sum_{j\neq i} K^{\varepsilon(i)*(j)}$, $\sum_{j\neq i} K^{\beta(i)*(j)}$	$i\varepsilon O$; $j\varepsilon I,0$	3.17a, 3.18, 3.19
OPTION 3	$H^{\varepsilon(i)\sigma(j)}$, $H^{\beta(i)\sigma(j)}$	$i\varepsilon I$; $j\varepsilon I,0$	3.16a, 3.17a
	$K^{\varepsilon(i)\sigma(j)}$, $K^{\beta(i)\sigma(j)}$	$i\varepsilon O$; $j\varepsilon I,0$	
OPTION 4	$H^{\varepsilon(i)\sigma(j)}$, $H^{\beta(i)\sigma(j)}$	$i\varepsilon I$; $J\varepsilon I$	
	$K^{\varepsilon(i)\sigma(j)}$, $K^{\beta(i)\sigma(j)}$	$i\varepsilon I,0$; $j\varepsilon O$	3.16b, 3.17b
	$\overline{H}^{\varepsilon(i)\sigma(j)}$, $\overline{H}^{\beta(i)\sigma(j)}$, $\sum_{j\neq i} K^{\varepsilon(i)*(j)}$, $\sum_{j\neq i} K^{\beta(i)*(j)}$	$i\varepsilon O$, $j\varepsilon I,0$	3.18, 3.19
OPTION 5	$K^{\varepsilon(i)\sigma(j)}$, $K^{\beta(i)\sigma(j)}$	$i\varepsilon I,0$; $j\varepsilon O$	
	$H^{\varepsilon(i)\sigma(j)}$, $H^{\beta(i)\sigma(j)}$	$i\varepsilon I$; $j\varepsilon I$	3.16b, 3.17b
	$K^{\varepsilon(i)\sigma(j)}$, $K^{\beta(i)\sigma(j)}$	$i\varepsilon O$; $j\varepsilon I,0$	
OPTION 6	$H^{\varepsilon(i)\sigma(j)}$, $H^{\beta(i)\sigma(j)}$	$i\varepsilon I,0$; $j\varepsilon I,0$	3.16, 3.17

CLOSED SYSTEM: options 1 and 6

SEMI-CLOSED SYSTEM: options 1,2,3,4, and 5.

The significance of the auto-projection option, described in section 1.1, is perhaps more evident now that the forecasting version of the model and its close relationship to the historical version have been outlined. Wherever it has been stated that a term must be estimated exogenously, the auto-projection development allows such a term to be estimated on a partly endogenous basis. This facility allows any input term to be transferred from the model run in the immediately preceding period to the forecasting run that follows. The term is, therefore, exogenous to the modelling period in question, but endogenous to the model run.

4. The Semi-Aggregate K-Date A.B.M.

In the totally aggregate version of the model that we have just described the use of a single deaths total ensures that the infant and non-infant populations are inextricably linked in the model through the single death rate that must be defined and through the nature of the accounting equations resulting from the available information. The conversion of this aggregate model into its "semi-aggregate" form, which considers the infant and non-infant populations more or less independently, is achieved in the historical case, therefore, by simply supplying two sets of death totals, one for each sub-population. This allows two sets of death rates to be defined, and also allows the accounting equations to be completely partitioned.

The structure of the resulting historical semi-aggregate model is illustrated in Figure D. As the non-infant and infant population sub-model are virtually identical we shall describe only the non-infant model equations to illustrate the differences between the aggregate and semi-aggregate versions, and leave readers the simple task of spelling out the infant equations for themselves:-

a. KNOWN INFORMATION. The following data is directly available from existing sources:-

$_K\varepsilon(i)*(*)$ Initial population

$_K\varepsilon(i)\sigma(j)$ Surviving non-infant migrants

FIGURE D: THE STRUCTURE OF THE HISTORICAL SEMI-AGGREGATE K-DATA A.B.M.

NON-INFANTS SECTION

Assemble known data

$_K\varepsilon(i)*(*)$

$_K\varepsilon(i)\sigma(j)$

$_K\varepsilon(*)\delta(i)$

Initial estimates of population at-risk of dying

$_K\hat{D}\varepsilon(*)i$ (4.1)

Calculate death rates

$d^{\varepsilon(*)i}$ (4.2)

Recalculate at-risk populations

$_K\hat{D}\varepsilon(*)i$ (4.6)

Calculate migrant deaths

$_K\varepsilon(i)\delta(j)$ $j\varepsilon I$ (4.3)

$_K\varepsilon(i)\delta(j)$ $j\varepsilon O$ (4.4)

Calculate non-migrant terms

$_K\varepsilon(i)\delta(i)$ (4.5)

$_K\varepsilon(i)\sigma(i)$ (3.12)

Convergence Tests

Not converged

Converged

Calculate final populations

$_K\varepsilon(*)\sigma(i)$

Calculate Population at-risk of giving birth and birth rates

$_K\hat{B}i*$ (4.6)

b^{i*} (3.3)

INFANTS SECTION

Assemble known data

$_K\beta(i)*(*)$

$_K\beta(*)\delta(i)$

Estimate surviving infant migrants

$_K\beta(i)\sigma(j)$ (3.1)

(as in non-infants section with β substituted for ε)

and the following must be estimated (exogenously) from existing sources (see Appendix B):-

$K^{\varepsilon(*)\delta(i)}$ Non-infant deaths

b. NON-INFANT POPULATION AT RISK OF DYING (INITIAL ESTIMATE). The best initial estimate is probably:-

$$\hat{K}^{D\varepsilon(*)i} = K^{\varepsilon(i)*(*)} + 0.5 \sum_{j \neq i} K^{\varepsilon(j)\sigma(i)}$$

$$- 0.5 \sum_{j \neq i} K^{\varepsilon(i)\sigma(j)} - 0.5\, K^{\varepsilon(*)\delta(i)} \tag{4.1}$$

c. NON-INFANT DEATH RATE.

$$d^{\varepsilon(*)i} = \frac{K^{\varepsilon(*)\delta(i)}}{\hat{K}^{D\varepsilon(*)i}} \tag{4.2}$$

d. NON-INFANT MIGRANT DEATHS. As in the aggregate model these minor flows are estimated in two parts, but now using the non-infant instead of the average death rate. For migrants dying in the main system:-

$$K^{\varepsilon(j)\delta(i)} = 0.5 \frac{d^{\varepsilon(*)i} \quad K^{\varepsilon(j)\sigma(i)}}{1-0.25d^{\varepsilon(*)i}} \quad i\varepsilon I;\ j\varepsilon I,O \tag{4.3}$$

and for migrants dying in external regions:-

$$K^{\varepsilon(j)\delta(i)} = 0.5 \frac{d^{\varepsilon(*)j} \quad K^{\varepsilon(j)\sigma(i)}}{1-0.25d^{\varepsilon(*)j}} \quad i\varepsilon O;\quad j\varepsilon I,O \tag{4.4}$$

e. NON-INFANT NON MIGRANTS. A revised accounting equation can now be used to calculate the non-migrant deaths:-

$$K^{\varepsilon(i)\delta(i)} = K^{\varepsilon(*)\delta(i)} - \sum_{j\neq i} K^{\varepsilon(j)\delta(i)} \qquad (4.5)$$

Non-migrant survivors are calculated using the same accounting equation as in the aggregate case (i.e.: equation 3.12).

Convergence tests are now applied, and if they are not satisfied the model goes through a cyclic procedure, recalculating the non-infant population at-risk of dying and subsequent flows as in the aggregate model.

f. FULL AT-RISK POPULATION. The version of the equation used here is:-

$$\hat{}_{K}D\varepsilon(*)i = K^{\varepsilon(i)\sigma(i)} + 0.5 \sum_{j\neq i} K^{\varepsilon(j)\sigma(i)} + 0.5\, K^{\varepsilon(i)\delta(i)} + 0.25 \sum_{j\neq i} K^{\varepsilon(j)\delta(i)} + 0.5 \sum_{j\neq i} K^{\varepsilon(i)\sigma(j)} + 0.25 \sum_{j\neq i} K^{\varepsilon(i)\delta(j)} \qquad (4.6)$$

Once convergence is achieved and the non-infant accounts sub-matrix completed, the final non-infant populations can be calculated. The population at-risk of giving birth is now defined, although, as in the aggregate historical case, it is redundant in modelling terms. It does not need to be calculated iteratively in this form of the model-

g. POPULATION AT-RISK OF GIVING BIRTH. By coincidence this is the same, under our present terms of reference.

as the population at-risk of dying given in equation (4.6). The birth rate is again defined using equation (3.3)as in the aggregate model.

In the historical case the two sub-models are totally independent except that the non-infant surviving migrant terms and initial populations may be used to estimate the infant surviving migrants. The structure of the two sub-models is also identical apart from this surviving migrant calculation. Indeed the model equations for the infant population may be defined simply by substituting a (β) label wherever an (ε) label occurs in the equations that we have just defined for the non-infant population, and also by dividing all the multiplicative weights on each term in the at-risk and minor flow equations by two to reflect the assumption that infants are likely to be exposed for only half the modelling period. It should perhaps be noted that non-migrant infant deaths now become major rather than minor flows as they can be estimated using an accounting equation in the semi-aggregate model as a result of the use of two deaths sub-totals.

The relationship between the historical and forecasting versions of the semi-aggregate model is very similar to that between the two versions of the aggregate model. Equally the difference between the semi-aggregate and aggregate versions of the historical accounts-based model are also evident in the forecasting case. In the forecasting version of the semi-aggregate model the infant and non-infant section are directly linked. The total births, required in the infants sub-model as the equivalent to the initial population, is produced after the non-infant accounts have been completed. In terms of their form the infant and non-infant sub-models in this version are identical, hence we have only shown the non-infant section in Figure E. Only one equation used in the forecasting case has not yet been defined. This is the equivalent of equation (3.21) in the aggregate model, and gives the deaths sub-total by rearranging equation (4.2):-

$$K^{\varepsilon(*)\delta(i)} = \hat{K}^{D\varepsilon(*)i} \quad d^{\varepsilon(*)i} \qquad (4.7)$$

FIGURE E: THE STRUCTURE OF THE SEMI-AGGREGATE FORECASTING A.B.M.

NON-INFANTS SECTION

Assemble Known data
$_K\varepsilon(i)*(*)$

Assemble estimated death rates
$d^{\varepsilon(*)i}$

Assemble estimated inputs for migrant calculations
(see Table 2 for options)

Calculate surviving non-infant migrants
(see Table 2 for equations)
$_K\varepsilon(i)\sigma(j)$

Initial Estimate of population at risk of dying
$_K\hat{D}\varepsilon(*)i$ (4.1)

Calculate deaths totals $\Leftarrow$ Recalculate at-risk populations $_K\hat{D}\varepsilon(*)i$ (4.6)
$_K\varepsilon(*)\delta(i)$ (4.7)

Calculate migrant deaths
$_K\varepsilon(i)\delta(j)$ $j\varepsilon I$ (4.3)
$_K\varepsilon(i)\delta(j)$ $j\varepsilon O$ (4.4)

Calculate non-migrant deaths
$_K\varepsilon(i)\delta(i)$ (4.5)
$_K\varepsilon(i)\sigma(i)$ (3.12)

Convergence tests $\Rightarrow$ Not converged

Converged

Calculate final populations
$_K\varepsilon(*)\sigma(i)$

Calculate population at-risk of giving birth and birth totals $\Leftarrow$ Input of estimated birth rates b^{i*}
$_K\hat{B}i*$ (4.16)
$_K\beta(i)*(*)$ (3.20)

INFANTS SECTION

Similar to non-infants section: β substituted for ε

All the remaining model equations are referred to in Figure E. All the migration estimation options, described for the aggregate forecasting model, also apply to this version as specified there. Birth and death information is still to be input in the form of rates, except that, in line with other changes, two death rates are now required in place of the original one.

5. The semi-aggregate M-data A.B.M.

We noted earlier, in section 1.4, that rarely outside the British case is migrant data directly available in the form required in the semi-aggregate and aggregate models in sections 3 and 4. In this section we shall set out a semi-aggregate model which uses an alternative migration data source and form - that of the number of migrations from the continuous registration system of recording. It must be remembered, however, that we are assuming no multiple migration takes place, and as a result we do not include any migration surplus terms in the model equations that follow.

The structure of the model is again illustrated diagrammatically (in Figure F). The model equations are as follows for the non-infants section of the model:-

a. KNOWN INFORMATION. The following data is directly available from existing sources:-

$K^{\varepsilon(i)*(*)}$	Initial population
$K^{\varepsilon(*)\sigma(i)}$	Final population

and the following must be estimated exogenously from other existing sources:-

$K^{\varepsilon(*)\delta(i)}$	Non-infant deaths
$M^{\varepsilon(i)*(j)}$	Non-infant migrants*

---

* Note that under our assumption of no surplus migration:-

$$M^{\varepsilon(i)*(j)} = K^{\varepsilon(i)*(j)} = K^{\varepsilon(i)\sigma(j)} + K^{\varepsilon(i)\delta(j)} \tag{5.0}$$

FIGURE F: THE STRUCTURE OF THE HISTORICAL SEMI-AGGREGATE M-DATA A.B.M.

NON-INFANT SECTION

Assemble known data

$K^{\varepsilon(i)*(*)}$

$K^{\varepsilon(*)\sigma(i)}$

$K^{\varepsilon(*)\delta(i)}$

$M^{\varepsilon(i)*(j)}$

↓

Initial estimate of population at-risk of dying

$\hat{K}^{D\varepsilon(*)i}$ (5.1)

↓

Calculate death rates ← Recalculate at-risk populations $\hat{K}^{D\varepsilon(*)i}$ (4.6)

$d^{\varepsilon(*)i}$ (4.2)

↓

Calculate surviving migrants

$K^{\varepsilon(i)\sigma(j)}$ $j\varepsilon I$ (5.2)

$K^{\varepsilon(i)\sigma(j)}$ $j\varepsilon O$ (5.3)

↓

Calculate migrant deaths

$K^{\varepsilon(i)\delta(j)}$ $j\varepsilon I$ (4.3)

$K^{\varepsilon(i)\delta(j)}$ $j\varepsilon O$ (4.4) or (5.4)

↓

Calculate non-migrants

$K^{\varepsilon(i)\delta(i)}$ (4.5)

$K^{\varepsilon(i)\sigma(i)}$ (3.12)

↓

Convergence tests → Not converged

↓

Converged

↓

Calculate population at-risk of giving birth and birth rates

$\hat{K}^{Bi*}$ (4.6)

b^{i*} (3.3)

↓

INFANTS SECTION

↓

Similar to non-infants section:

β substituted for ε

b. NON-INFANT POPULATION AT RISK OF DYING (INITIAL ESTIMATE). We cannot use the initial estimate from the K-data version of the model as some of the terms are unknown at this point. In its place we use an average non-infant population:-

$$\hat{{}_{K}}^{L\varepsilon(*)i} = 0.5\ ({}_{K}^{\varepsilon(i)*(*)} + {}_{K}^{\varepsilon(*)\sigma(i)}) \qquad (5.1)$$

c. NON-INFANT DEATH RATE. The same equation (4.2) is used here as in the K-data version of the model.

d. NON-INFANT SURVIVING MIGRANTS. These flows are estimated in two sets as the migrant deaths have been in previous models. For migrants surviving in the main system:-

$${}_{K}^{\varepsilon(j)\sigma(i)} = \frac{{}_{M}^{\varepsilon(j)*(i)}}{1 + \left[\dfrac{0.5\ d^{\varepsilon(*)i}}{1-0.25d^{\varepsilon(*)i}}\right]} \qquad i\varepsilon I; \quad j\varepsilon I,0 \qquad (5.2)$$

and for migrants surviving in external regions:-

$${}_{K}^{\varepsilon(j)\sigma(i)} = \frac{{}_{M}^{\varepsilon(j)*(i)}}{1 + \left[\dfrac{0.5\ d^{\varepsilon(*)j}}{1-0.25\ d^{\varepsilon(*)j}}\right]} \qquad i\varepsilon 0; \quad j\varepsilon I \qquad (5.3)$$

e. NON INFANT DEATHS. The terms may either be calculated from the equation :-

$${}_{K}^{\varepsilon(j)\delta(j)} = {}_{M}^{\varepsilon(j)*(i)} - {}_{K}^{\varepsilon(j)\sigma(i)} \qquad (5.4)$$

given the relationship in equation (5.0), or they can be calculated as usual in two sets as minor flows using equations (4.3) and (4.4) from the semi-aggregate K-data version of the model.

4. NON-INFANT NON-MIGRANTS. Again the equations used in the K-data version (4.5 and 3.12) are appropriate.

Convergence tests are applied at this point and the iterative process is adopted, recalculating the population at-risk of dying (using equation 4.6) and the subsequent terms accordingly, until convergence is achieved.

The redundant births information is calculated as in the historical semi-aggregate K-data model. The infant section of the model, which is identical in form to the non-infant section just described, may then be calculated independently with the appropriate adjustments to the non-infant model equations that were indicated in section 4.

No forecasting version of this particular model has been formulated. We feel that the forecasting version of the section 4 model may be used satisfactorily for all semi-aggregate work.

6. The Rogers G-matrix forecasting model

This model is described in an accounting context in Rees and Wilson (1977, Chapter 6) along with other transition rate models, and an example of its use is to be found in Rees (1976). It has been included in the programmes as a simple alternative to the forecasting versions of the accounts-based model.

The model, as specified in the programmes, applies a matrix of "growth" rates to a set of base populations to produce a set of forecasted populations. In matrix terms we may define the model as:-

$$\underline{w}\,(t+1) = \underline{G}.\,\underline{w}\,(t) \tag{6.1}$$

where $\underline{w}$ (t) is the column vector of regional populations at time t; $\underline{G}$ is the growth matrix whose elements are defined (in accounting terms) by the equation:-

$$G^{ij} = \frac{K^{\varepsilon(j)\sigma(i)} + K^{\beta(j)\sigma(i)}}{K^{\varepsilon(j)*(*)}} \quad i\varepsilon I,O; \quad j\varepsilon I,O \tag{6.2}$$

and $\underline{w}$ (t+1) is the column vector of populations produced for the end of the forecasting period (time t+1).

In the programmes that follow there is a choice between using a closed or semi-closed system with the model. The model we have just specified refers to the closed-system case. A modified version is used in the semi-closed case which incorporates the contribution of migration into each main system region from external regions by a vector of surviving migrants. The model is then redefined as:-

$$\underline{w}\ (t+1) = \underline{G} \cdot \underline{w}(t) + \underline{m} \tag{6.3}$$

where $\underline{m}$ is the inmigrant vector, and equation (6.2) is more accurately respecified to give the elements of the growth matrix in this version as :-

$$G^{ij} = \frac{K^{\varepsilon(j)\sigma(i)} + K^{\beta(j)\sigma(i)}}{K^{\varepsilon(j)*(*)}} \quad i\varepsilon I,O; \quad j\varepsilon I \tag{6.4}$$

In this forecasting model the only estimated inputs required are the growth matrix in the closed-system case, and the inmigrant vector in addition in the semi-closed case. The auto-projection option, allowing old values of these inputs to be transferred between modelling periods instead of exogenously estimating them, also applies to this particular model. Each element of the surviving inmigrant vector ($\underline{I}$) is defined in accounting terms as:-

$$m^{j} = \sum_{i}(K^{\varepsilon(i)\sigma(j)} + K^{\beta(i)\sigma(j)}) \quad i\varepsilon O; \quad j\varepsilon I \tag{6.5}$$

PART III: THE MULTIPOP2 PROGRAMME

In this third part of the paper we aim to provide a detailed description of one of the three programmes that have been developed to implement some of the accounts-based models described in Rees and Jenkins (1977). We shall concentrate on the MULTIPOP2 programme as the most direct descendant of the SDAT programme of Rees and Wilson (1974), and probably the most relevant to the majority of users. Limited details of the other two programmes appear in Appendix C, and further information of interest will be supplied to readers who wish to use either of the two programmes.

This part of the paper may be considered as a user's manual for the MULTIPOP2 programme. In section 7 we describe the programme's structure by discussing the function of each of its subroutines. This section is cross-referenced with section 9, where the inputs and outputs for two examples are set out in as much detail as a paper of this sort will allow. In section 8 we describe the necessary particulars for creating the input card-decks or files required to run examples with the programme. Input listings for the various types of model run are given to illustrate the points made earlier in the section on programme control and data inputs.

7. The form of the programme

MULTIPOP2 is a computer programme which provides the means of running the aggregate historical and forecasting versions of the accounts-based model described in section 3. The programme has been specified as a series of subroutines controlled by a main segment. Particular modelling tasks are allocated to each subroutine as indicated in Figure G. The Rogers G-matrix forecasting model described in secrion 6 and provided as an alternative to the forecasting version of the accounts-based model, is specified as one of these subroutines.

We shall descrube the function of each of the subroutines in turn (in the order in which they appear) to illustrate the structure of the programme, and also as a means of relating the programme to the discussion of the model in part II of the paper. Although each subroutine has a particular task, it may be used in more than one version of the model since, as section 3 has shown, there is much in common, in terms of model equations, to both the historical and forecasting accounts-based models. This multiple use of subroutines, together with the savings resulting from the similarity of many of the non-infant and infant model equations, has, to some extent, reduced the length of the programme text, but we make no claims as to having defined the most computationally efficient programme. We do hope, however, that we have produced a programme that, with the notes that follow, will be a clear and effective one.

We shall refer to an example to illuminate several points about the programme. This example, which is intended to give readers some feel for the sort of results that they may expect to produce with the programme, is derived from information in a report by Rees (1977b). It relates to a four-region system, with main system regions of East Anglia, the South-East and the rest of Britain, and a rest-of-the world external region. The example is set out in full in section 9 for easy reference.

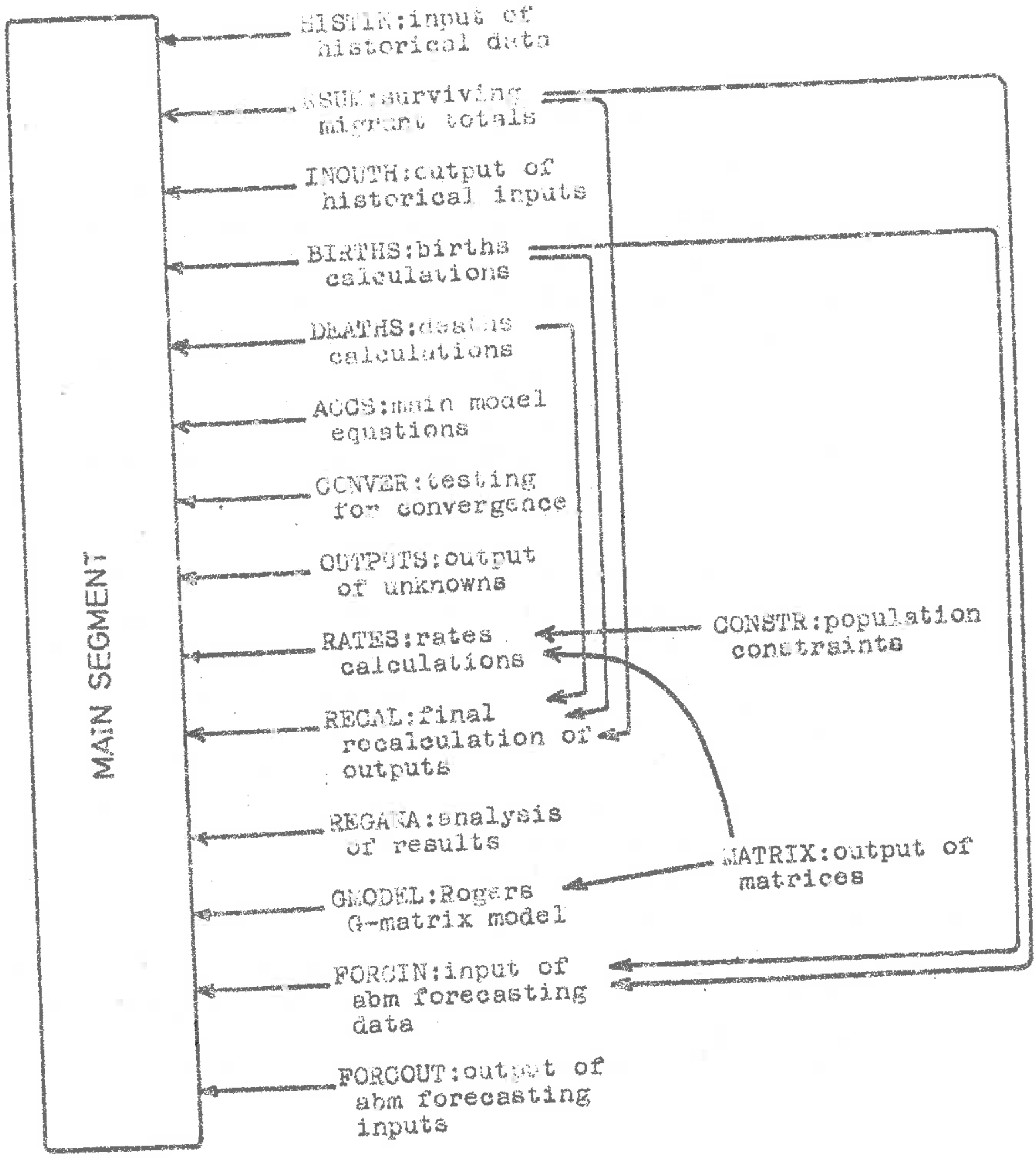


FIGURE 6 : SUBROUTINE FUNCTIONS IN THE MULTIPOP2 PROGRAMME

The primary task of the main segment of the programme is to control the subroutines by calling them into operation in the correct sequence when required. It does perform a series of other tasks, however, and the first of these is to read in the first set of programme control cards.* These cards contain the basic information on the nature of the example to be run, such as the number of regions in the system, the number of historical and forecasting periods for which the model is to be run, which version of the a.b.m. is to be used initially and so on. Having read in the control parameters, the regional labels and two title cards, the main segment prints out a title sequence summarising the information. This output is given for our four-region example in Figure J1. Another task of the main segment is to ensure that parameters and variables are set to the appropriate value before each run or iteration of the model is performed. Before calling any subroutines the initial population of each region, data common to all versions of the accounts-based model, is also read in via the main segment.

Subroutine HISTIN reads in all the other inputs required by the historical version of the a.b.m., and also calculates the surviving infant migrant terms if they have not already been estimated exogenously. Subroutine KSUM calculates various surviving migrant totals from information read in, or estimated in HISTIN (in the historical case) or FORCIN (in the forecasting case). These migrant totals are printed out, in the historical case, together with all input data and any terms estimated in HISTIN, in subroutine INOUTH. Figure J2 gives the output produced by this subroutine for our four-region example. The migrant totals given are defined as follows. The (II) immigration figures are the sum of migration into each region from all other main system regions :-

$$\sum_{j \neq i} K^{\alpha(j)\sigma(i)} \quad i\varepsilon I,O, \quad j\varepsilon I$$

---

* We shall use the term "cards" throughout, even though inputs to the programme may also be from a data file.

The (OI) immigration figures are the sum of migration into each main system region from all external regions:-

$$\sum_{j\neq i} K^{\alpha(j)\sigma(i)} \quad i\varepsilon I, \quad j\varepsilon O$$

This term will be zero in a closed system case as there are no external regions, and in the semi-closed system case it will also be set to zero for all external regions. Finally the (*I) immigration figures will be the total migration into main system regions from all other regions (that is the sum of the II and OI terms):-

$$\sum_{j\neq i} K^{\alpha(j)\sigma(i)} \quad i\varepsilon I; \quad j\varepsilon I,O$$

In a closed system case this term will be the same as the (II) immigration figure in every region. In the semi-closed system case this term will again be set to zero for all external regions. The set of emigration totals is defined in exactly the same manner as those for immigration.

Subroutine BIRTHS performs all calculations connected with births in both the historical and forecasting versions of the model. In the historical case the populations at risk of giving birth in each region and the birth rates are estimated, and in the forecasting case the at-risk populations and total births are produced. Equation (3.15) is used to produce both the initial and subsequent estimates of the at-risk populations, the unknown terms being set to zero in the case of the initial estimate (thereby producing equation 3.2). Parallel to subroutine BIRTHS is subroutine DEATHS which has exactly the same functions and form with respect to the deaths calculations.

Subroutine ACCS performs all the minor flow and accounting equations which, as we have already noted, are common to both versions of the model. The convergence tests on the elements of the accounts matrix are carried out in subroutine CONVER. Subroutine OUTPUTS prints out all the terms that have been estimated in the model's iterative cycle - that is, all the terms in the accounts matrix that are unknown, together with the non-surviving migrant totals, the final populations and the terms calculated in subroutines BIRTHS and DEATHS. This subroutine is again used in both versions of the model. Since the unknowns differ between versions the output produced also differs. Figure J3 gives the output produced by the subroutine for a historical run with our four-region example, and Figure J13 for a typical forecasting case. A parameter (NPRINT) is read in at the start of each run of the models which indicates whether the user wants the results of all model iterations to be printed out, or only the values produced in the final iteration (that is, the one on which convergence is achieved). If only the final results are required, as in our example, the OUTPUTS subroutine is by-passed for all but the last iteration.

The next subroutine, RATES, called once the accounts-based model has converged, has several functions. The first of these is to output the estimated accounts matrix together with its row and column totals (see Figure J4). In the historical case the subroutine CONSTR will then be called if the ICONS parameter, read in at the start of the programme, indicates that the user requires this estimated accounts matrix to be constrained to the observed row and column totals, using the methods described in section 1.5. Subroutine CONSTR prints out various pieces of information which may be of interest to the user. Firstly it gives the value of four statistics which compare the initial estimated and observed constraint end-of-period population terms. These are the index of dissimilarity, the coefficient of correlation (r), the coefficient of determination (r^2) and the sum of squared deviations. Obviously these statistics must be interpreted with caution, particularly where, as in our example, only a small number of regions are being used. The second set of output information relates to the balancing factor routine itself. The final values of the balancing factors, as given by equations (1.4) and (1.5), are printed out for each row and column of the accounts matrix, together with the remaining discrepancy between the revised row and column totals produced by these balancing factors and the observed totals. If the balancing factor routine has not converged after the maximum number of iterations allowed by the user (through the input parameter ITER) the results produced in the last iteration are adopted for use in the subsequent modelling operations. In this case the output from this subroutine may prove useful in indicating how near to convergence the routine actually is. Figure J5 gives the output produced by this subroutine for our example. Control is now returned to subroutine RATES and the revised accounts matrix is printed out (see Figure J6).

The next task of subroutine RATES is to calculate and output the transition rates matrix (H), the admission rates matrix (A), and the growth rates matrix (G) from the accounts matrix (see Figures J7, J8 and J9). The transition rates matrix, potentially required to estimate migration inputs in the forecasting version of the accounts-based model (as indicated in section 3.2), is produced by dividing each element of the final accounts matrix by its row total. For a closed system we can define a general equation for this as:-

$$H^{\alpha(i)\omega(j)} = \frac{K^{\alpha(i)\omega(j)}}{K^{\alpha(i)*(*)}} \quad i\varepsilon I,O; \quad j\varepsilon I,O \tag{7.1}$$

As we noted earlier, in the semi-closed system case all accounts flows totally within the external regions are not modelled, and are set to zero in the accounts matrix. The row totals for the external regions, therefore, give the total migrants from each external region (for infants and non-infants) to the main system regions. As a result the transition rates matrix defined and output in the semi-closed case is really a mixture of two types of rates:-

$$H^{\alpha(i)\omega(j)} = \frac{K^{\alpha(i)\omega(j)}}{K^{\alpha(i)*(*)}} \quad i\varepsilon I; \quad j\varepsilon I,O \tag{7.2}$$

and

$$\overline{H}^{\alpha(i)\omega(j)} = \frac{K^{\alpha(i)\omega(j)}}{\sum\limits_{j\neq i} K^{\alpha(i)\omega(j)}} \quad i\varepsilon O; \quad j\varepsilon I,O \tag{7.3}$$

Two options of the forecasting model have been specified, however, which use the rates in this mixed form to produce the migration inputs. Two further options ignore the $\overline{H}$ rates and use only the

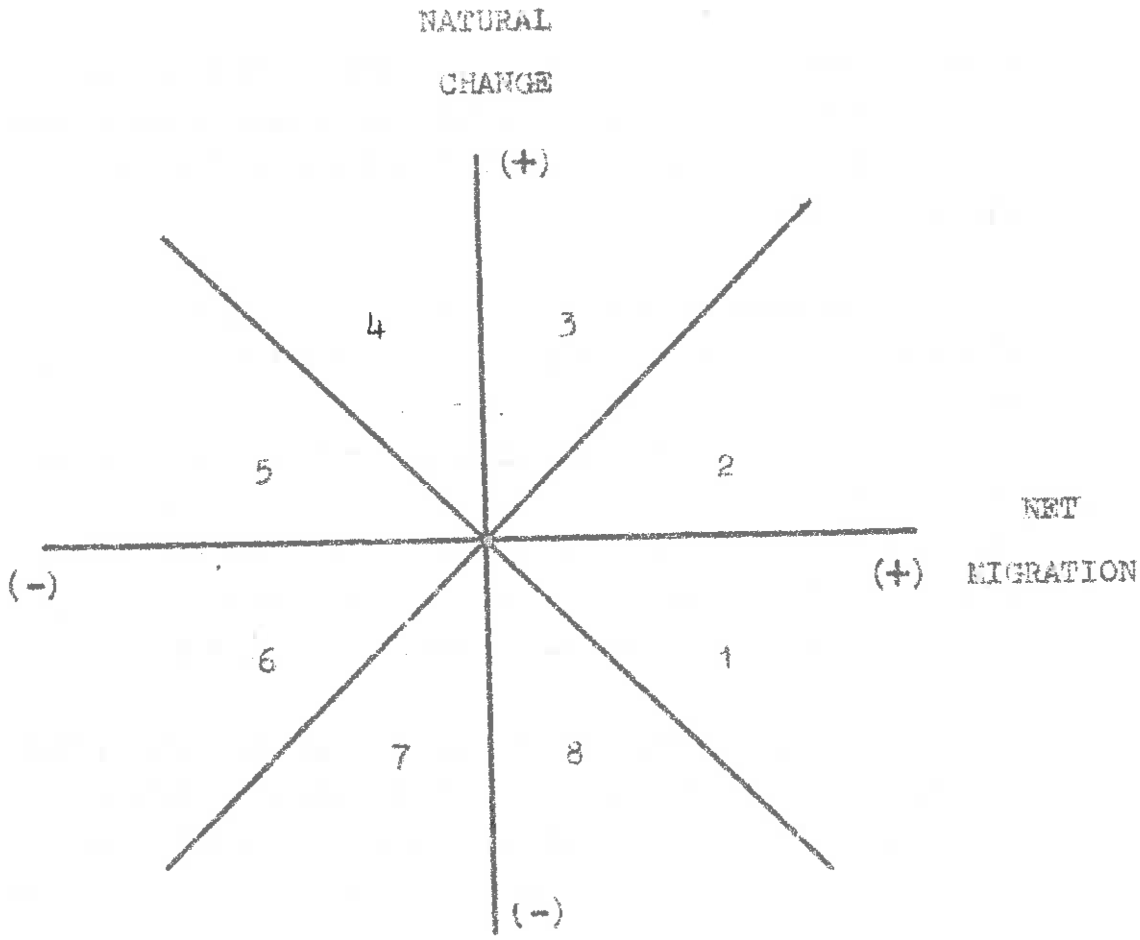
H rates defined in equation (7.2). A fifth option uses no transition rates at all, and the sixth option (which may be used in the closed system case) uses the H rates as defined in equation (7.1).

The admission rates matrix, which is not required in any of the forecasting models in this paper, is defined by dividing each element of the final accounts matrix by its column total. We leave users to define the equations for themselves, and remind them that this matrix will also combine two types of rate in the semi-closed system case. The growth rates matrix has already been defined in section 6. As we noted that it may be used in what we have termed the Rogers G-matrix forecasting model.

The subroutine RATES itself calls another subroutine, MATRIX, to output the matrices that we have just defined. MATRIX divides these matrices into blocks of eleven columns' width, wherever necessary, as larger matrices cannot be printed as they stand owing to the limited width of the output medium.

If, in the historical case, the balancing factor routine has been used to produce a revised accounts matrix consistent with the observed end-of-period populations subroutine RECAL is called from the main segment. This subroutine recalculates various terms using the revised, rather than the originally-estimated, accounts matrix. These terms are the migrant totals (infant and non-infant, surviving and non-surviving, immigrating and outmigrating), and the birth and death rates and at-risk populations. These are output as in Figure J10.

Subroutine REGAMA analyses the results produced by the accounts-based models for each period. It classifies each main region into one of eight population categories according to the relationship between the level of natural change and net migration. These eight categories are shown in Figure H together with the equivalent in the original paper by Webb (1963). It also compares the percentage distribution of population betwen regions in the main system at the start and end of each modelling period, as well as giving the percentage change in each region's population over the period. Finally the total population of the main system as a whole



GROUPS 1-4 : POPULATION INCREASE

1 NET IMMIGRATION exceeds NATURAL LOSS(D)

2 NET IMMIGRATION exceeds NATURAL INCREASE(C)

3 NATURAL INCREASE exceeds NET IMMIGRATION(B)

4 NATURAL INCREASE exceeds NET OUTMIGRATION(A)

GROUPS 5-8 : POPULATION DECREASE

5 NET OUTMIGRATION exceeds NATURAL INCREASE(H)

6 NET OUTMIGRATION exceeds NATURAL LOSS(G)

7 NATURAL LOSS exceeds NET OUTMIGRATION(F)

8 NATURAL LOSS exceeds NET IMMIGRATION(E)

FIGURE H : POPULATION CLASSIFICATION BY THE LEVEL OF NET MIGRATION AND NATURAL CHANGE

(The letters after each group are Webb's(1963)categories) is given for both times. Figure J11 shows the output produced by this subroutine for our four region example.

The Rogers G-matrix forecasting model, described in section 6 is run by the subroutine GMODEL, and once called by the main segment operates independently for as many forecasting periods as required as input statements, model equations and outputs are all contained within the subroutine. This may be a useful point to note that, as we have specified the programme, this model may only be used in conjunction with a prior base-run of the historical accounts-based model. The results, as output by this subroutine, of the Rogers G-matrix model run for our four region example are given in Figure K.

Subroutine FORCIN is the forecasting model equivalent of subroutine HISTIN. As well as reading in those new inputs required in each forecasting period, as specified by the set of control cards for the forecasting series (which are also read in by this subroutine), it estimates the non-infant and infant surviving migrant matrices wherever necessary. As we indicated in section 3.2 the infant migrant terms must be calculated iteratively since they are estimated from the births totals. This means that this input subroutine becomes part of the iterative cycle unlike the historical case where it was not. Subroutines KSUM and BIRTHS are called directly from FORCIN as a result of its role in estimating the migrant terms. The final subroutine in the programme is FORCOUT, the forecasting equivalent of INOUTH. An example of the output produced by this subroutine is given in Figure J12. Exactly the same sequence of subroutines is used in the forecasting version of the accounts-based model as in the historical, once subroutines FORCIN AND FORCOUT have been called in place of HISTIN and INOUTH. The only difference is that subroutines CONSTR and RECAL, which relate to the constraining of the estimated accounts matrix to observed population totals in the historical model to produce a revised accounts matrix, are redundant in the forecasting case as no observed end-of-period populations will be available. Consequently in this version of the model these two subroutines are ignored.

8. Input Instructions for the MULTIPOP2 Programme.

The aim of this section is to describe in detail the series of input instructions required by the MULTIPOP2 programme so that users may run whichever model options they find most useful. Examples of sets of input decks are given to illustrate a selection of the alternatives.

We may summarize the input requirements of the programme as:-

(a) a series of programme control cards indicating the general details of the example to be run, and specifying the nature of any historical models required;

(b) data for the base-run of the model, which may be, in this case, a historical or forecasting version of the accounts-based model;

(c) another set of programme control cards for any forecasting model to be run;

(d) data for any subsequent modelling periods that is to differ from that in the immediately preceding period, whether that period be a base-run or an earlier forecasting period.

We begin by describing all the control cards in turn, and by noting briefly the function of each control parameter, the possible values it may take, and any limitations that need to be borne in mind with respect to other control parameters and data inputs. We then set out the various data inputs required in line with the chosen control parameters. Finally in this section we give a series of examples of full input decks for various types of modelling operation. This will show, in a summary form, which control and data cards are required for each particular type of run, and the order in which they must be input to the programme.

8.1 "Control Cards"

After each programme control card name the format in which the data is to be specified is given in computing terms. In the first column in each card-description the column numbers in which

(continued on page 60)

CARD A — Format 6I3

1-3	N	The total number of regions in the example. A maximum of 50 is available in this programme.
4-6	NIS	The 'code' number of the first external region. If a closed system is being used then NIS must be set to N+1. In a semi-closed system the regions should be ordered such that all the external regions are the last listed. Any number of external regions may be used within the limits of the total number of regions, so that NIS may take any value from 2 to N in this case. If there is only one external region, for example, NIS will take the same value as N; if there are two NIS will equal N-1 and so on.
7-9	NHP	The code specifying the mode of the first model version to be run:- NHP < 0 means a historical a.b.m. NHP ≥ 0 means a forecasting a.b.m. This first run is termed the "base run", since all subsequent model runs use it for reference purposes.
10-12	MFOR	The number of "auto-projection" periods to follow any base run (historical or forecasting). A maximum of 50 has been specified.
13-15	NITMAX	The maximum number of iterations to be allowed in the a.b.m. estimations before the run is ended on the assumption that convergence will not be achieved - usually no more than 10 are needed.
16-18	NPRINT	The code specifying whether the results of all iterations are to be output or only the final results. If only the final results are required NPRINT must be set to zero. All other values produce output for each model iteration, for each modelling period. Generally only the final results will be of interest.

CARD B Format I3

This card is only required if a historical base run of the a.b.m. is to be carried out (that is, if NHP < 0 on card A).

1-3	MBS	This parameter indicates whether infant migrants in the historical version of the model are to be read into the programme, having been calculated externally (if so MBS should be set to 1), or whether they are to be calculated internally using equation (3.1), in which case MBS should be set to 2.

Note that if the second option is selected in a semi-closed system case the initial populations and birth totals must be supplued for all regions so that the estimations can be made. These pieces of data are not required for any other calculations, and once used in this estimation are discarded.

CARD C Format I4, I2

This card is not required if only a historical run is being carried out (that is if NHP < 0 OR MFOR = 0 on card A). The information on this card is used to label the output of forecasting model woek.

1-4	NYEAR	The year in which forecasts are to commence.
5-6	NPER	The length of the forecasting period(s) in years. This parameter is assumed to be constant over the series of modelling periods.

CARD D   Format 2I3

This card is only required if there is to be a historical run of the accounts-based model (that is when NHP < 0 on card A).

1-3	ICONS	The ICONS parameter should be set to zero if the balancing factor subroutine is not required, and to a positive integer in format I3 if it is required.
4-6	ITER	The maximum number of iterations to be allowed in the iterative balancing factor routine before a halt is called if convergence has not been achieved. The number of iterations required by the routine will vary from problem to problem depending on the number of regions, the size of the errors in the populations and so on. We advise a value of at least 150 for this parameter. If convergence is not achieved within the limit, the revised accounts matrix is produced using the values of the balancing factors after the last iteration to be completed.

CARD SET E   Format 20A4

Each region may be given a label which is used in printing out the results. Each region has been allocated a unit of four alphanumeric characters for its label. Regional names should be right-justified - that is, if less than four characters are being used in the label spaces should be inserted to the left of the characters to make up a unit of four.

The regional labels should be in the same order as the regional data inputs:-

1-4	WORDS(1)	Region label for zone 1
5-8	WORDS(2)	Region label for zone 2
	...etc...	

Each card has eighty columns and so a maximum of twenty region labels may be accommodated on any one card. Where over twenty regions are being used the labels should be continued on a new card.

Example for our four-region system in section 9:-

--EA--SE--RB--RTW

(- indicates a space).

CARD SET F

Two cards have been provided so that users may input their own title to head the whole set of results that are output. Any alphanumeric characters may be used, with a maximum of 80 characters on each card.

CARD G Format I3

This card is not required if only a single historical run of the accounts-based model is to be carried out (that is if NHP < 0 AND MFOR = 0 on card A).

1-3	MODEL	A parameter specifying which form of forecasting model is required in the "auto-projection" sequence:-
		0 indicates the Rogers G-matrix forecasting model (see section 6)
		1 indicates the forecasting version of the a.b.m. (see section 3.2).

CARD H Format 50I1

This card is only required when the Rogers G-matrix forecasting model is to be used, as indicated on card G by a zero value for the MODEL parameter.

The function of this card is to control the reading in of new data for the series of auto-projection forecasting periods. There will be MFOR such forecasting periods (see card A), and this card will carry MFOR columns of parameters, one for each period. The first parameter, in column 1, will specify for which inputs to the model, if any, new values are to be read in for the first period, and for which the "old" values from the immediately preceding model run are to be adopted. For this period the preceding model run will be either a historical or a forecasting base-run of the A.B.M. The second parameter (in column 2) will indicate the same information but for the second period, and the preceding model run now becomes the first auto-projection run of the G-matrix model. The third parameter relates to the third period and so on.

The choice of parameters is set out in Table 3, together with two examples of the card's format. Both examples relate to a semi-closed system with five forecasting periods using the Rogers G-matrix model. Users must remember that they may not change from a closed to a semi-closed system or vice-versa within the forecasting series.

The maximum number of forecasting periods allowed is 50.

CARD I Format I3

This card is only required if a forecasting base-run of the accounts-based model is to be used (that is if NHP ≥ 0 on card A).

1-3	MCLOSE	This parameter specifies the form in which the data (from which the migrant terms are to be produced), is to be input into the forecasting a.b.m. The various options are described in detail in section 3.2, and summarised in Table 2. To recap, in a closed system MCLOSE may be either 1 or 6, and in a semi-closed system the parameter may take any value from 1 to 5.

CARD J Format 50I1

This card is only required when the forecasting version of the accounts-based model is to be used in auto-projection, as indicated by a value of 1 for the MODEL parameter on card G.

It has the same form as card H, and the same function as card I, that is, a parameter relating to each of the forecasting periods in turn must be entered in each column indicating the form in which the data, from which the migrant terms are produced in this version of the model, is to be input. The options are summarised in Table 2 in section 3.2.

The comments made in relation to card H and I also apply to this card, and will not be repeated. Once again a maximum of 50 forecasting periods is allowed.

CARD SET K Format 50I1
(on each of three cards)

As for card J, this set of three cards is only required when the forecasting version of the accounts-based model is to be used in auto-projection mode.

The form of each of the three cards is similar to that of card H, with one column for each forecasting period, up to a maximum of 50. Each of the cards performs the same function, with respect to the forecasting version of the a.b.m., as that of card H in relation to the Rogers G-matrix model. There are three cards instead of one, however, as the a.b.m. has considerably more input options available. The first of the three cards, MDATA (1, INDIC), controls the data required to produce the migration terms in each forecasting period; the second, MDATA (2, INDIC), the birth rates; and the third, MDATA (3, INDIC), the death rates.

Tables 2 and 4, and card J, should be referred to in deciding upon the values to be placed on the first of these cards. Having decided, and entered on card J, the form in which the data to produce the migrant terms is to be input to the model, users must then decide whether the value of each input term (as summarised in Table 2) is to be the "old" one, from the preceding period, or a "new" one which needs to be read in. Table 4 indicates for each MCLOSE option all the possible combinations of "old" and "new" data terms that are available in any period, and the MDATA (1.INDIC) parameter value associated with each.

The choice of parameters for the second and third cards in the set is far more limited and, as a result, more straightforward. On both cards a zero indicates that "old"rates are to be used in a particular period, and a one that "new" rates are to be read in.

TABLE 3: THE CHOICE OF DATA-INPUT CONTROL PARAMETERS FOR THE ROGERS' G-MATRIX FORECASTING MODEL

Value of GDATA (I)
(I = number of forecasting period)

Closed System Case

0	"Old" G-matrix to be used
1	"New" G-matrix to be read in

Semi-Closed System Case

2	"Old" G-matrix and surviving inmigrant vector(s)
3	"New" G-matrix and surviving inmigrant vector(s)
4	"Old" G-matrix and "New" surviving inmigrant vector(s)
5	"New" G-matrix and "Old" surviving inmigrant vector(s)

Examples

22422	Two forecasting periods with a constant G-matrix and migrant vectors; new migrant vectors then to be read in (with G-matrix still constant), and three forecasting periods with this new migrant vectors held constant. The G-matrix held constant over the whole period.
33424	Two forecasting periods in both of which new G-matrices and migrant vectors are read in; new migrant vectors then to be read in and used for two forecasting periods, and a further [illegible] of new vectors to be read in for the fifth period; the G-matrix read in for the second period held constant to the end of the forecasts.

MCLOSE (INDIC)		MDATA (1,INDIC)							
		0	1	2	3	4	5	6	7
	Input Form (see Table 3)								
1	Flows	x	✓	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
2	Transition rates	x	✓	x	✓	n.a.	n.a.	n.a.	n.a.
	Inmigrant totals	x	✓	✓	x				
3	Transition rates	x	✓	x	✓	n.a.	n.a.	n.a.	n.a.
	Inmigrant flows	x	✓	✓	x				
4	Transition rates	x	✓	x	✓	x	✓	✓	x
	Inmigrant totals	x	✓	✓	x	x	✓	x	✓
	Outmigrant flows	x	✓	x	x	✓	x	✓	✓
5	Transition rates	x	✓	x	✓	x	✓	✓	x
	Inmigrant flows	x	✓	✓	x	x	✓	x	✓
	Outmigrant flows	x	✓	x	x	✓	x	✓	✓
6	Transition rates	x	✓	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

x = "OLD VALUES from preceding model run

✓ = "NEW" VALUES to be read in

n.a. = this option does not apply

TABLE 4: THE CHOICE OF DATA-INPUT CONTROL PARAMETERS FOR THE FORECASTING VERSION OF THE ACCOUNTS-BASED MODEL THE MIGRANT TERMS

(Note: INDIC is the label for the forecasting period number) each piece of data must be entered are given as a check on the format statement. In the second column in the description the name given to each input parameter in the programme is given. In the third column the possible values of each parameter are discussed and any limitations on the values they may take are noted. Note that on all the control cards all input parameter values should be <u>right-justified</u> (that is, where the number of columns allocated to any particular value exceeds the number of digits in the actual value spaces should be inserted to the left of the input digits to make up the extra columns).

8.2 Data Cards

All data should be input to the programme in free format. The only convention that must be followed with such a format is that each piece of data must be separated from the next by one or more spaces. Since the data inputs obviously vary according to the type of model selected for any particular period, we feel the best way to approach the question of input form and order is to describe a series of general modelling cases, and the data associated with each.

In a "<u>simple historical run</u>", in which the historical accounts-based model is run for a single period, only one set of data cards is required. This would be as given in Table 5. Note that two possible input schedules have been set out there. In thefirst (where MBS = 1 on control card B) surviving infant migrants are to be calculated exogenously and are consequently considered as an input to the programme, while in the second (where MBS = 2) these migrant terms are calculated within the programme. The second alternative requires, as a result, that figures for the initial populations and total births of <u>all</u> regions be supplied, rather than simply those for the main system regions only. The only other information that may be needed in a simple historical run is a set of observed row and column constraints, where the user has requested the application of the balancing factor routine to the estimated accounts matrix (through control card D). These constraints are the totals to which the accounts matrix is to be adjusted, and they should be read in immediately following the data presented in Table 5. The row constraints should be entered first, followed by the column constraints. Where a new value for any constraint is not to be input, and the

	MBS=1	MBS=2
Initial Populations (KOLDP - $K^{\varepsilon(i)*(*)}$)	NIS-1 values, with zero dummies for any external regions.	N values (no dummies)
E.S. Migration Matrix (K(I,J) - $K^{\varepsilon(i)\sigma(j)}$)	N x N, with zero diagonals and zero dummies for totally external flows.	
Births Totals (KTB - $K^{\beta(i)*(*)}$)	NIS-1 values, with zero dummies for any external regions.	N values (no dummies)
Deaths Totals (KTD - $K^{*(*)\delta(i)}$)	NIS-1 values, with zero dummies for any external regions.	
B.S. Migration Matrix (K(N+I,J) - $K^{\beta(i)\sigma(j)}$)	NxN matrix, with zero diagonals and zero dummies for totally external flows	NOT REQUIRED

Note: The migration matrices should be entered row by row, rather than column by column.

TABLE 5: DATA INPUTS FOR THE HISTORICAL ACCOUNTS - BASED MODEL

existing row and column totals of the accounts matrix (as estimated by the model) are to be accepted, a value of zero should be entered for that particular constraint. A new constraint may be input for any of the individual row or column totals of the accounts matrix, whether initial or final population, birth or deaths total, or migrant total (in the semi-closed system case). It is vital that the row and column constraints to be used in the balancing factor routine are equal when summed, or else the routine will not converge satisfactorily. It is left to users to ensure that this condition is satisfied before the constraints are input, but an error message is printed out if it has not been satisfied, and the unadjusted accounts matrix is then used in any further modelling.

In a "simple forecasting run" (a single-period run of the forecasting a.b.m.) the data inputs required would be as given in Table 6 for the appropriate option selected, preceded by a set of initial populations entered as in the simple historical run example (for the MBS = 1 case). The population constraint option does not apply to the forecasting case, and so no further information is necessary in this type of run.

In the "auto-projection (a.b.m. forecasting version) case" two sets of data are required. The first relates to the base-run of the model and would be exactly as for the "simple" cases just described, in either historical or forecasting mode. The second set of data, generally separated from the first in the input card by a card or cards controlling the auto-projection exercise, relates to subsequent forecasting runs of the a.b.m. As we have already indicated, only that data which is to differ from the previous period's needs to be input. There is complete flexibility in that none, some or all of the inputs can be set at the value of those in the previous run if required. This choice is made through control cards J and K described earlier. The data cards for the auto-projection periods must be ordered as in the "simple" forecasting case (see Table 6) BUT ONLY THOSE NEW INPUTS WHICH ARE REQUESTED BY THE CONTROL CARDS MUST BE ENTERED ON THE DATA CARDS FOR EACH PERIOD IN TURN.

TABLE 6: DATA INPUTS FOR THE SIX OPTIONS OF THE FORECASTING VERSION OF THE ACCOUNTS BASED-MODEL.

MCLOSE = 1

E.S. Migration Matrix $K^{\varepsilon(i)\sigma(j)}$ $i\varepsilon I,O$; $j\varepsilon I,O$

NxN matrix, with zero diagonals and zero dummies for totally external flows

Birth Rates b^{i*} $i\varepsilon I,O$

NIS-1 values with zero dummies for any external regions

Death Rates d^{*i} $i\varepsilon I,O$

NIS-1 values with zero dummies for any external regions

B.S. Migration Matrix $K^{\beta(i)\sigma(j)}$ $i\varepsilon I,O$; $j\varepsilon I,O$

(as for E.S. Migration Matrix)

MCLOSE = 2

E.S. Transition Rates $H^{\varepsilon(i)\sigma(j)}$ $i\varepsilon I,O$; $j\varepsilon I,O$

NxN matrix, with zero diagonals and zero dummies for totally external flows

(E) Inmigrant Total $\sum_{j\neq i} K^{\varepsilon(i)*(j)}$ $i\varepsilon O$; $j\varepsilon I,O$

For each external region

Birth Rates
Death Rates } as in MCLOSE = 1

B.S. Transition Rates $H^{\beta(i)\sigma(j)}$ $i\varepsilon I,O$; $j\varepsilon I,O$

(as for E.S. transition rates)

(B) Inmigrant Total $\sum_{j\neq i} K^{\beta(i)*(j)}$ $i\varepsilon O$; $j\varepsilon I,O$

For each external region

MCLOSE = 3

E.S. Transition Rates $H^{\varepsilon(i)\sigma(j)}$ $i\varepsilon I$; $j\varepsilon I,O$

(NIS-1) x N matrix with zero "diagonals"

(E) Inmigrant Vectors $K^{\varepsilon(i)\sigma(j)}$ $i\varepsilon O$; $j\varepsilon I,O$

For each external region, a vector of flows, with zero dummies for flows to other external regions

Birth Rates
Death Rates } as in MCLOSE = 1

B.S. Transition Rates $H^{\beta(i)\sigma(j)}$ $i\varepsilon I$; $j\varepsilon I,O$

(as for E.S. transition rates)

(B) Inmigrant Vectors $K^{\beta(i)\sigma(j)}$ $i\varepsilon O$; $j\varepsilon I,O$

(as for (E) inmigrant vectors)

TABLE 6 (continued)

MCLOSE = 4

E.S.Transition Rates	$H^{\varepsilon(i)\sigma(j)}$	$i\varepsilon I,O$; $j\varepsilon I$
N x (NIS-1) matrix, with zero "diagonals"		
(E) Immigrant Totals (as in MCLOSE = 2)	$\sum_{j\neq i} K^{\varepsilon(i)*(j)}$	$i\varepsilon O$; $j\varepsilon I,O$
(E) Outmigrant Vectors	$K^{\varepsilon(i)\sigma(j)}$	$i\varepsilon I,O$; $j\varepsilon O$
For each external region, a vector of flows, with zero dummies for flows from other external regions.		
Birth Rates } as for MCLOSE = 1		
Death Rates		
B.S. Transition Rates (as for E.S. transition rates).	$H^{\beta(i)\sigma(j)}$	$i\varepsilon I,O$; $j\varepsilon I$
(B) Immigrant Totals (as in MCLOSE = 2)	$\sum_{j\neq i} K^{\beta(i)*(j)}$	$i\varepsilon O$; $j\varepsilon I,O$
(b) Outmigrant Vectors (as for (E) outmigrant vectors)	$K^{\beta(i)\sigma(j)}$	$i\varepsilon I,O$; $j\varepsilon O$

MCLOSE = 5

E.S. Transition Rates	$H^{\varepsilon(i)\sigma(j)}$	$i\varepsilon I$; $j\varepsilon I$
(NIS - 1) x (NIS - 1) matrix, with zero diagonals		
(E) Immigrant Vectors (as in MCLOSE = 3)	$K^{\varepsilon(i)\sigma(j)}$	$i\varepsilon O$; $j\varepsilon I,O$
(E) Outmigrant Vectors (as in MCLOSE = 4)	$K^{\varepsilon(i)\sigma(j)}$	$i\varepsilon I,O$; $j\varepsilon O$
Birth Rates } as for MCLOSE = 1		
Death Rates		
B.S. Transition Rates (as for E.S. transition rates)	$H^{\beta(i)\sigma(j)}$	$i\varepsilon I$; $j\varepsilon I$
(B) Immigrant Vectors (as for (E) immigrant vectors)	$K^{\beta(i)\sigma(j)}$	$i\varepsilon O$; $j\varepsilon I,O$
(B) Outmigrant Vectors (as for (E) outmigrant vectors)	$K^{\beta(i)\sigma(j)}$	$i\varepsilon I,O$; $j\varepsilon O$

TABLE 6 (continued)

MCLOSE = 6

E.S. Transition Rates $H^{\varepsilon(i)\sigma(j)}$ $i\varepsilon I,0;$ $j\varepsilon I,0$

N x N matrix, with zero diagonals

Birth Rates
Death Rates } as for MCLOSE = 1

B.S. Transition Rates $H^{\beta(i)\sigma(j)}$ $i\varepsilon I,0;$ $j\varepsilon I,0$

(as for E.S. transition rates)

NOTES

(1) The migration and rates matrices should be entered row by row, rather than column by column.

(2) The inmigrant vectors should be entered destination by destination for each external region in turn. The outmigrant vectors should be entered origin by origin for each external destination region in turn.

In the "auto-projection (Rogers G-matrix model) case" two sets of data are again required. The first set again relates to the base-run of the a.b.m., and data would be input exactly as indicated in the previous example. The second set of data, again separated from the first set by a control card, relates to the subsequent forecasting runs with the Rogers G-matrix model. For each period in turn, data must be input for ONLY THE NEW TERMS REQUESTED THROUGH CONTROL CARD H in the form and order specified in Table 7. The elements of the surviving inmigrant vector in the semi-closed case are defined as :-

$$K^{*(R)\sigma(j)} = \sum_{i\varepsilon O} (K^{\varepsilon(i)\sigma(j)} + K^{\beta(i)\sigma(j)}) \quad j\varepsilon I \qquad (8.1)$$

TABLE 7: DATA INPUTS FOR THE ROGERS G-MATRIX FORECASTING MODEL

CLOSED SYSTEM	SEMI-CLOSED SYSTEM
G-MATRIX	G-MATRIX
N x N, matrix with zero dummies for totally external flows	(NIS - 1) x N matrix with zero dummies for totally external flows
	Surviving inmigrant vector Row vector of NIS-1 values.

8.3 Example Input Decks/Files

Inputs to the programme may be either through the medium of cards or in a data file. Instructions for running jobs at Leeds by both means are given in Appendix A. The actual records are identical, however, in both cases and so in the examples that follow no reference will be made to any particular type of input medium.

In this section we shall set out four complete examples of the inputs required to run jobs for each of the general modelling cases outlined in the previous section. These will illustrate which particular control cards are required with particular types of modelling exercise, and the order in which they and the data cards must be arranged.

An example of the inputs for a "simple historical run" is given in Table 8. Control card A (line zero) specifies that the example relates to four regions, of which the fourth is external to the main system. A single historical run of the a.b.m.is to be made, in which the maximum number of model iterations to be allowed is 10. Only the final results are to be printed out. Control card B (line 1) specifies that the surviving infant migrant matrix is to be estimated exogenously and read in to the programme later in the input list. Control card C is not required in a simple historical run and so the next record in the input list is control card D (line 2). This specifies that population constraints are to be applied to the estimated accounts matrix. A maximum of 200 iterations has been allocated for convergence to be achieved in this routine. Card E (line 3) holds the regional labels for the four zones in the example. We again remind users, at this point, of the rule of right-justification of all entries, as illustrated in these first four lines, that must be followed. Card set F (lines 4 and 5) carries the title sequence that will head the output. Lines 6 to 18 inclusive carry the data required by the historical a.b.m. in the order given in Table 5. The initial populations are given on line 6; lines 7-10 carry the non-infant surviving migrants matrix, line 11 the births totals and line 12 those for deaths; lines 13-16 give the infant surviving migrants matrix; and finally lines 17 and 18 give the population constraints for the start and end of the modelling period respectively that the accounts matrix must satisfy. No further control or data cards are required in this example. We shall refer to the function of line 19 later in the section.

An example of the inputs for a "simple forecasting run" is given in Table 9. Control card A (line zero) differs in this example from the simple historical run example only in parameter NHP, which now indicates that a forecasting base run of the a.b.m. is required. Control Card B . is not required in this type of run. The next record in the input list is control card C (line 1) which specifies that forecasts begin in 1971 and that a five-year forecasting period is to be adopted. This information is used solely for labelling purposes. Control card D is not required, and card E (line 2), carrying the regional labels, follows. Card set F (line 3 and 4)

"SIMPLE HISTORICAL RUN" : a single period run of the historical version of the accounts-based model.

CONTROL CARDS   A,B,D,E,F   LINES 0-5

DATA CARDS   LINES 6-18

```
0    4  4,-1   0 10   0
1    1
2    3200
3    EA   SE   RB RTW
4     A FOUR REGION SEMI-CLOSED SYSTEM EXAMPLE USING MULTIPOP2
5  A SINGLE HISTORICAL RUN OF THE ACCOUNTS-BASED MODEL FOR 1966-71
6 1560535 16650328 34948572 0
7 0 51397 53941 34315
8 108034 0 566655 678122
9 83321 565593 0 716485
10 51060 516188 466291 0
11 131591 1385042 3007207 0
12 92514 955116 2090751 0
13 0 4202 3659 2331
14 6519 0 33211 44326
15 8847 30178 0 46971
16 3425 24608 31473 0
17 1570938 16698754 34704488 1121280 0 0 0 63631
18 1681754 16988102 35369209 1463655 0 0 0 41830
19 000
```

TABLE 8 : EXAMPLE OF AN INPUT LISTING FOR A "SIMPLE HISTORICAL RUN"

"SIMPLE FORECASTING RUN" : a single period run of the forecasting version of the accounts-based model .

CONTROL CARDS I	A,C,E,F	LINES 0-4
DATA CARD I	LINE 5	
CONTROL CARD II	I	LINE 6
DATA CARDS II ·	LINES 7-18	

```
 0    4  4  1  0 18  0
 1 1971 5
 2   EA  SE  RB RTW
 3   A FOUR REGION SEMI-CLOSED SYSTEM EXAMPLE USING MULTPOP2
 4 A SINGLE FORECASTING RUN OF THE ACCOUNTS-BASED MODEL FOR 1971-76
 5 1681754 16988102 35369209 0
 6   5
 7 0 0.0369913 0.0381339
 8 0.0057026 0 0.0334445
 9 0.0021291 0.0164514 0
10 52988 609769 541057 0
11 40001 701508 745684 0
12 0.08492 0.08586 0.08890 0
13 0.05725 0.05682 0.05922 0
14 0 0.0359450 0.0307470
15 0.0041356 0 0.0235574
16 0.0026407 0.0102078 0
17 3563 29139 36607 0
18 2705 45709 49260 0
19 000
```

TABLE 9 : EXAMPLE OF AN INPUT LISTING FOR A "SIMPLE FORECASTING RUN"

again carries a title sequence. Line 5 gives the base populations from which the forecasts are to commence. This is separated from the rest of the data by control card I (line 6), which specifies that the migrant terms required in the forecasting version of the a.b.m. are to be estimated/input according to the methods of option 5 in Table 2. Control cards G,H,J and K are not required. Lines 7-18 inclusive hold the remaining data required by the forecasting model in the order given in Table 6 (under MCLOSE = 5). Lines 7 - 9 give the exist-survive transition rates matrix; line 10 the non-infant inmigrant terms, line 11 the non-infant outmigrant terms; lines 12 and 13 carry the birth and death rates respectively; and lines 14-18 give the same terms as lines 7-11 except that they now refer to infants.

Moving on to the more complex types of case, we give an example of the inputs for an "auto-projection (forecasting a.b.m.) run" in Table 10. Control card A (line zero) here is virtually identical to that of the simple historical run in Table 8. The only change required is in the parameter indicating the number of auto-projection periods to run (MFOR). In our example we have specified three such periods to follow the historical base-run of the a.b.m. Control card B (line 1) is as in the "simple" example, indicating the same choice of treatment for the infant migrant terms. Control card C (line 2) is now required as there are forecasts to follow the historical modelling. Control card D (line 3) relates to this historical base run only, and in this example indicates that no population constraints are to be applied to the accounts. Cards E and F (lines 4,5 and 6) are as described earlier. Following these control cards comes the data relating to the historical model run (lines 7-17). This is specified exactly as it was in the "simple" historical case. Note that as no population constraints are to be applied the equivalent of lines 17 and 18 in Table 8 are absent in the Table 10 example. All the lines from 18 onwards in Table 10 refer to the auto-projection runs of the forecasting version of the a.b.m. Control card G (line 18) indeed indicated that it is the accounts-based model, rather than the Rogers G-matrix model, that is to be used in the forecasting exercise. Control card J (line 19) specifies that in the first and second forecasting periods the migrant terms are to be estimated/input using the methods of option 4 in Table 2, and of option 2 for the third forecasting period. Control card set K

AUTO-PROJECTION(Forecasting a.b.m.)RUN : a base-run of the historical a.b.m., followed by a series of auto-projection runs with the forecasting version of the accounts-based model.

CONTROL CARDS I	A,B,C,D,E,F	LINES 0-6
DATA CARDS I	LINES 7-17	Historical data
CONTROL CARDS II	G,J,K	LINES 18-22
DATA CARDS II	LINES 23-27	First forecasting period
	LINES 28-30	Second forecasting period

```
0    4  4 -1  3 10  0
1    1
2 1971 5
3    0  0
4    EA  SE  RB RTW
5    A FOUR REGION SEMI-CLOSED SYSTEM EXAMPLE USING MULTIPOP2
6   1966-71 ACCOUNTS FOLLOWED BY 1971-86 FORECASTS WITH A.B.M.
7 1560535 16650328 34948572 0
8 0 51397 53941 34315
9 198034 0 566655 678122
10 83321 565593 0 716485
11 51060 516188 466291 0
12 131591 1385042 3007207 0
13 92514 255116 2090751 0
14 0 4202 3659 2331
15 5519 0 33211 40326
16 4897 30178 0 46971
17 3425 24608 31473 0
18    1
19 442
20 740
21 100
22 010
23 100000
24 35000 650000 700000 0
25 0.08492 0.08586 0.08890 0
26 65000
27 2600 42500 47500 0
28 32500 675000 700000 0
29 0.05600 0.05575 0.05870 0
30 2500 40000 48000 0
31 000
```

TABLE 10 : EXAMPLE OF AN INPUT LISTING FOR AN AUTO-PROJECTION (FORECASTING A.B.M.)RUN--HISTORICAL BASE-RUN

(lines 20-22) specifies for which particular variables new values are to be read in in each period. In the first forecasting period the new data inputs expected are given by the first column of lines 20-22, together with the first column of line 19. Thus, consulting Table 4, we may see that, with a value of 4, for the MCLOSE parameter for period one (line 19) and a value of 7 for the MDATA parameter for the migrant terms for that period (line 20), new inmigrant totals and outmigrant flows are to be read in, while the same transition rates produced in the preceding modelling period (i.e. the historical base-run) are to be adopted. In addition new birth rates are to be input for this first forecasting period, as indicated by the one in column one of line 21, while the deaths rate from the previous period are to be used, as specified by a zero in column one of line 22. The order in which the three new sets of variables are to be read in is given in Table 6 under the appropriate MCLOSE value (which is four in this case). The new data for this first forecasting period are entered in the input list immediately after the control cards on lines 19-22. In this example the data on lines 23 to 27 inclusive relates to the first period. That on lines 28 to 30 relates to the second forecasting period, and has been requested in the second column of the control cards on lines 19-22. No new data has been requested for the third forecasting period, and so no data needs to be supplied after line 30.

Our final input example, given in Table 11, refers to the "auto-projection (Rogers G-matrix model) case". Lines zero to 17 inclusive are identical in format and order to the auto-projection example just described using the forecasting a.b.m. This is because both are using a historical accounts-based model as a base run for the forecasting exercise. Here, however, the Rogers G-matrix forecasting model is requested on control card G (line 18). Control card H (line 19) performs the same function in relation to the Rogers model as control card set K did for the forecasting version of the a.b.m. in the last example. Thus in the first column both a new G-matrix and a new surviving inmigrant vector are requested for the first forecasting period (see Table 3). These are entered in lines 20 to 23 inclusive, following the control card that specifies that they are required. The data on line 24 is the new surviving inmigrant vector requested for the second forecasting period in column two of control card H. No new data is

AUTO-PROJECTION(Rogers G-matrix model)RUN : a base-run of the historical a.b.m., followed by a series of auto-projection runs with the Rogers G-matrix forecasting model .

CONTROL CARDS I	A,B,C,D,E,F	LINES 0-6
DATA CARDS I	LINES 7-17	Historical data
CONTROL CARDS II	G,H	LINES 18+19
DATA CARDS II	LINES 20-23	First forecasting period
	LINE 24	Second forecasting period

```
 0    4  4 -1  3 10  0
 1    1
 2 1971 5
 3    0  0
 4   EA  SE  RB RIW
 5       A FOUR REGION SEMI-CLOSED SYSTEM EXAMPLE USING MULTIPOP2
 6 1966-71 ACCOUNTS FOLLOWED BY 1971-86 FORECASTS WITH ROGERS MODEL
 7 1560535 16650328 34948572 0
 8 0 51397 53941 34315
 9 108034 0 566655 678122
10 83321 565593 0 716485
11 51060 516188 466291 0
12 131591 1385042 3007207 0
13 92514 955116 2090751 0
14 0 4202 3659 2331
15 6519 0 33211 44326
16 8887 30178 0 46971
17 3425 24608 31473 0
18    0
19 342
20 0.9181890 0.0060457 0.0023579 0.0456328
21 0.0400022 0.9392748 0.0173359 0.5155556
22 0.0407094 0.0353984 0.9836332 0.4661359
23 60000 600000 550000
24 62500 620000 525000
25 000
```

TABLE 11 : EXAMPLE OF AN INPUT LISTING FOR AN AUTO-PROJECTION (ROGERS G-MATRIX MODEL)RUN--HISTORICAL BASE-RUN

required for the third forecasting period as the third column on control card H has been set at a value of two.

In the two auto-projection examples described we have used the historical version of the accounts-based model in the base-runs. The programme also allows the forecasting a.b.m. to be used for this purpose if so desired. If the forecasting version of the accounts-based model is also to be used in the auto-projection series following the base-run, an appropriate input list could be produced by using exactly the same form and order of control and data cards as in the "simple forecasting run" (except that the MFOR parameter on control card A would need to be changed from zero to the number of auto-projection periods required), followed by control cards G, J and K and the data they specify, exactly as in lines 18 to 30 of the auto-projection (forecasting a.b.m.) run in Table 10. If the Rogers G-matrix model is to be used in the auto-projection series following a forecasting base-run of the accounts-based model, control cards G and H, and the data they specify, should follow revised control and data cards of the "simple forecasting run".

We now have six basic types of programme run that can be carried out. These are summarised in Table 12, together with the control cards that each type of run requires and the value that certain key input parameters must take.

In each of the four examples in this section, the final line of the input list has been composed of three noughts in the first three columns. The function of this line is to indicate to the programme that there is no information to be read in. A message stating that the job has been completed is then written out and the programme run is halted. There is, however, no reason why only one modelling run should be performed in any one job. Consequently we could make one large input listing out of the four examples in this section by combining them into one card deck or file. Lines 0 to 18 from Table 8, would be followed by lines 0 to 18 of Table 9, followed by lines 0 to 30 of Table 10, followed by lines 0 to 24 of Table 11 and completed then by the three noughts and the four stars on the last two lines. In other words blocks of complete control and data input sets can be repeated in a single input deck or file for as many cases as there are examples to be run if users think they stand to gain by this facility. Each deck or file must be

TYPE OF RUN	CONTROL CARDS REQUIRED AND THEIR ORDER	"INDICATIVE PARAMETERS"	
Simple: a single run of the accounts-based model.			
"SIMPLE HISTORICAL"	A;B;D;E;F	NHP < 0	MFOR = 0
"SIMPLE FORECASTING"	A;C;E;F;I	NHP ≥ 0	MFOR = 0
AUTO-PROJECTION: a base run of the historical or forecasting version of the a.b.m., followed by one or more forecasting runs			
HISTORICAL BASE-RUN	A;B;C;D;E;F	NHP < 0	MFOR > 0
FORECASTING A.B.M.	G;J;K	MODEL = 1	
FORECASTING BASE RUN	A;C;E;F;I	NHP ≥ 0	MFOR > 0
FORECASTING A.B.M.	G;J;K	MODEL = 1	
HISTORICAL BASE RUN	A;B;C;D;E;F;	NHP < 0	MFOR > 0
ROGERS G-MATRIX	G;H	MODEL = 0	
FORECASTING BASE-RUN	A;C;E;F;I	NHP ≥ 0	MFOR > 0
ROGERS G-MATRIX	G;H	MODEL = 0	

TABLE 12: A SUMMARY OF SIX TYPES OF PROGRAMME RUN

terminated, however, by the three noughts of the programme terminator card, and whatever system terminators are required by particular machines (e.g: four stars in the case of the GEORGE system at Leeds).

We should point out to readers that, in the forecasting examples that have been given in this section, the assumptions about the future population development of the regional system under consideration, and the data inputs for the forecasting periods, are not intended to represent our objective view. They are purely illustrative of the form and order of inputs in particular cases.

9. Input and Output - Two Examples

The aim of this final section is to give readers some idea of the nature of the output that they may expect to obtain from the MULTIPOP2 programme. It is not possible to reproduce the output in exactly its original form in a paper of this sort, but we have attempted to preserve the essential features of a typical output sequence. As our aim is largely illustrative, we have provided only one example of any section of output which appears in every period for which the models are run. Generally in the original, each section of output begins on a new page, but, to save space, we have combined two sections on one page in a couple of cases.

In our first example we have run the historical accounts-based model for a single inter-censal period (1966-71) for a four-region system centred on East Anglia, with data from Rees (1977b), followed by forecasts for a further five year period to 1976 using the forecasting version of the a.b.m. A new set of birth rates are read into the programme for use in the forecasts, otherwise all rates are assumed constant at the 1966-71 levels estimated in the historical base-run. The input file, FOURREGION1, containing all the necessary programme control information and data, is reproduced in Table 13, and the output produced in the various parts of Figure J.

The second example uses the same four-region system as the first. In this case, however, the 1966-71 historical run of the accounts-based model is followed by forecasts for the 1976 populations produced with the Rogers G-matrix model. The input file for this example FOURREGION2.

is reproduced in Table 14. The output of the Rogers model run in given in Figure K, but that from the historical base-run is not repeated as it is identical to that of the previous example in Figure J.

AUTO-PROJECTION(Forecasting a.b.m.)RUN : a base-run of the historical a.b.m. for 1966-71, followed by a run of the forecasting version of the accounts-based model for 1971-76.

CONTROL CARDS I	A,B,C,D,E,F	LINES 0-6
DATA CARDS I	LINES 7-19	Historical data
CONTROL CARDS II	G,J,K	LINES 20-24
DATA CARD II	LINE 25	"New" forecasting data

```
 0    4  4 -1  1 10  0
 1    1
 2 1971 5
 3    3200
 4    EA  SE  RB RTW
 5       A FOUR REGION SEMI-CLOSED SYSTEM EXAMPLE USING MULTIPOP2
 6 1966-71 ACCOUNTS FOLLOWED BY 71-76 FORECAST WITH NEW BIRTH RATES
 7 1560535 16650328 34948578 0
 8 0 51397 53941 34315
 9 108034 0 566655 678122
10 83321 565593 0 716485
11 51060 516188 466291 0
12 131591 1385042 3007207 0
13 92514 955116 2090751 0
14 0 4202 3659 2331
15 6519 0 33211 44326
16 8887 30178 0 46971
17 3425 24608 31473 0
18 1570938 16698754 34704488 1121280 0 0 0 63631
19 1681754 16988102 35369209 1463655 0 0 0 41630
20    1
21 5
22 0
23 1
24 0
25 0.072947 0.069418 0.072238 0.0
26 000
```

TABLE 13 : INPUT FILE FOURREGION1

FIGURE J: OUTPUT FROM A TYPICAL AUTO-PROJECTION (Forecasting a.b.m.) RUN OF MULTIPOP2, USING DATA FILE FOURREGION1

1. Main title and information summary
2. Historical period: input data
3. Historical period: estimates of unknowns
4. Historical period: estimated accounts matrix
5. Historical period: population constraints routine
6. Historical period: revised accounts matrix
7. Historical period: final transition-rates matrix
8. Historical period: final admission-rates matrix
9. Historical period: final growth-rates matrix
10. Historical period: final estimates of unknowns
11. Historical period: analysis of results
12. Forecasting period: "input" data
13. Forecasting period: estimates of unknowns

```
************************************************************************************************************************
*****************       A FOUR REGION SEMI-CLOSED SYSTEM EXAMPLE USING MULTIPOP2                              ***********
*****************   1966-71 ACCOUNTS FOLLOWED BY 71-76 FORECAST WITH NEW BIRTH RATES                          ***********
************************************************************************************************************************
```

BASE RUN OF THE MODEL USING HISTORICAL A.B.M.

WITH 1 AUTOMATIC PROJECTION PERIODS TO FOLLOW

FORECASTS COMMENCE IN 1971 WITH PERIODS OF 5 YEARS DURATION

THE ANALYSIS RELATES TO A SYSTEM OF 4 REGIONS,OF WHICH 1 EXTERNAL TO MAIN SYSTEM

ONLY THE FINAL RESULTS ARE TO BE PRINTED

THE MAXIMUM NUMBER OF ITERATIONS TO BE USED IN THE MODEL SET AT 10

FIGURE J1

FULLY AGGREGATE K ACCOUNTS BASED MODEL
****

INPUT DATA
****

	INITIAL POPULATION	BIRTHS	DEATHS
EA	1560935.	131391.	92514.
SE	10650328.	1383042.	955116.
RB	34946572.	3007207.	2096751.
RTW	0.	0.	0.

(A)MIGRATION+SURVIVAL TOTALS

	IMMIGRATION			EMIGRATION		
	II	OI	*I	II	IO	I*
EA	191355.	51060.	242415.	105338.	34315.	139653.
SE	616990.	516188.	1133178.	674689.	678122.	1352811.
RB	620596.	466291.	1086887.	648914.	716485.	1365399.
RTW	1426722.	0.	0.	1033539.	0.	0.

(B)MIGRATION+SURVIVAL TOTALS

	IMMIGRATION			EMIGRATION		
	II	OI	*I	II	IO	I*
EA	15406.	3425.	18831.	7861.	2331.	10192.
SE	34380.	24608.	58988.	39730.	44326.	84056.
RB	36870.	31473.	68343.	39065.	46971.	86036.
RTW	93028.	0.	0.	59506.	0.	0.

EXISTENCE AND MIGRATION MATRIX

EA	0.	51397.	53941.	34315.
SE	108034.	0.	566655.	678122.
RB	83321.	565593.	0.	716485.
RTW	51060.	516188.	466291.	0.

BIRTH AND MIGRATION MATRIX

EA	0.	4202.	3659.	2331.
SE	6519.	0.	33211.	44326.
RB	8887.	30178.	0.	46971.
RTW	3425.	24608.	31473.	0.

FIGURE J2

INFORMATION CALCULATED BY THE MODEL ON ITERATION NUMBER 4
****

****CONVERGENCE ACHIEVED ON THIS ITERATION****

	BIRTH		DEATH		MIGRATION +DEATH			
	AT RISK	RATE	AT RISK	RATE	(B)IMMIG	(E)IMMIG	(B)EMIG	(E)EMIG
EA	1569696.60	0.08383	1636749.36	0.05652	267.99	6940.21	148.11	4092.34
SE	16077000.20	0.08615	16753471.12	0.05761	846.76	32768.29	1224.78	39749.67
RB	33802009.81	0.08897	35279213.93	0.05926	1020.11	32690.41	1260.78	40293.66
RTW	0.00	0.00000	0.00	0.00000	1370.57	42142.85	871.78	30415.09

	(E)NON-MOVER TERMS		(B)NON-MOVER TERMS		FINAL POPULATION
	K(I,I)	K(I,N+I)	K(N+I,I)	K(N+I,N+I)	
EA	1335016.89	81774.77	117728.86	3522.03	1713989.754
SE	14373965.25	883784.08	1262042.35	37718.87	16828193.601
RB	31572362.45	1970517.09	2833386.83	86523.39	35560979.078
RTW	0.00	0.00	0.00	0.00	0.000

EXISTENCE AND DEATH MATRIX

EA	81774.7693	1486.2553	1622.3888	983.6939
SE	3096.9659	883782.0831	17043.3389	19609.3626
RB	2388.5285	16355.3435	1970517.0877	21349.7907
RTW	1463.7158	14926.6912	14024.6809	0.0000

BIRTH AND DEATH MATRIX

EA	3524.0308	60.3189	54.6154	33.1732
SE	92.7739	37718.8658	495.2178	636.2911
RB	126.4736	437.1993	86523.3947	701.1039
RTW	48.7422	353.2430	449.7759	0.0000

FIGURE J3

*** ACCOUNTS MATRIX ***

ROWS 1- 4 REFER TO THOSE ALIVE IN EACH REGION AT THE START OF THE PERIOD
ROWS 5- 8 REFER TO THOSE BORN IN EACH REGION DURING THE PERIOD
COLUMNS 1- 4 REFER TO THOSE SURVIVING TO THE END OF THE PERIOD IN EACH REGION
COLUMNS 5- 8 REFER TO THOSE DYING DURING THE PERIOD IN EACH REGION
THE LAST ROW GIVES THE COLUMN TOTALS
THE LAST COLUMN GIVES THE ROW TOTALS

ROW NUMBER BLOCK OF COLUMNS NUMBER 1 COLUMN NUMBERS 1 TO 9

ROW NUMBER	1	2	3	4	5	6	7	8	9
1	1335015.	51397.	53941.	34315.	81775.	1486.	1622.	984.	1560535.
2	708034.	14373965.	566655.	678122.	3097.	883782.	17043.	19609.	16650327.
3	83321.	565593.	31572362.	716485.	2389.	16355.	1970517.	21530.	34948572.
4	51060.	516188.	466291.	0.	1464.	14927.	14025.	0.	1063955.
5	117729.	4202.	3659.	2331.	3522.	60.	55.	33.	131591.
6	6519.	1262042.	33211.	44326.	93.	37719.	496.	636.	1385042.
7	8887.	30178.	2833387.	46971.	126.	433.	86523.	701.	3007236.
8	3425.	24608.	31473.	0.	49.	353.	470.	0.	60378.
9	1713990.	16828193.	35560979.	1522550.	92515.	953115.	2090751.	43513.	58807606.

FIGURE J4

RECALCULATION OF ACCOUNTS MATRIX TO OBTAIN CONSISTENCY WITH OBSERVED POPULATION CONSTRAINTS
****
(USING DOUBLY-CONSTRAINED BALANCING FACTOR ROUTINE)

POPULATION TOTALS

	START-OF-PERIOD		END-OF-PERIOD	
	MODEL	CONSTRAINT	MODEL	CONSTRAINT
EA	1560535.	1570938.	1713990.	1681754.
SE	16050327.	16698794.	16823193.	16985102.
GB	34948572.	34704488.	35560979.	35369209.

STATISTICS TO COMPARE OBSERVED AND ESTIMATED END-OF-PERIOD POPULATIONS
****

INDEX OF DISSIMILARITY = 0.3326

COEFFICIENT OF CORRELATION = 0.35833

COEFFICIENT OF DETERMINATION = 0.12840

SUM OF SQUARED DEVIATIONS = 63385780876.998

CONVERGENCE CRITERIA ACHIEVED AFTER 157 ITERATIONS

	BALANCING FACTORS		ERRORS IN TOTALS (ESTIMATED MINUS GIVEN CONSTRAINTS)	
	ROW(A)	COLUMN(B)	ROW	COLUMN
1	0.999927	1.000073	0.00	0.00
2	0.999927	1.000073	0.05	-0.00
3	0.999927	1.000073	-0.05	0.00
4	0.999927	1.000073	0.00	0.00
5	0.999927	1.000073	0.00	-0.00
6	0.999927	1.000073	0.00	-0.00
7	0.999927	1.000073	-0.00	0.00
8	0.999927	1.000073	0.00	-0.00

FIGURE J5

*** REVISED ACCOUNTS MATRIX ***

ROWS 1- 4 REFER TO THOSE ALIVE IN EACH REGION AT THE START OF THE PERIOD
ROWS 5- 8 REFER TO THOSE BORN IN EACH REGION DURING THE PERIOD
COLUMNS 1- 4 REFER TO THOSE SURVIVING TO THE END OF THE PERIOD IN EACH REGION
COLUMNS 5- 8 REFER TO THOSE DYING DURING THE PERIOD IN EACH REGION
THE LAST ROW GIVES THE COLUMN TOTALS
THE LAST COLUMN GIVES THE ROW TOTALS

ROW NUMBER   BLOCK OF COLUMNS NUMBER 1   COLUMN NUMBERS 1 TO 9

1	1329316.	58079.	59659.	36762.	82398.	1665.	1805.	1054.	1570936.
2	95805.	14475727.	558549.	647447.	2731.	682739.	16912.	18725.	16698753.
3	74516.	573944.	31358252.	687296.	2162.	16460.	1969125.	20733.	34704488.
4	48169.	552546.	488536.	0.	1397.	15847.	14784.	0.	1121279.
5	116061.	4725.	4027.	2485.	3531.	67.	61.	35.	131592.
6	5762.	1265827.	34603.	42150.	83.	37522.	450.	605.	1385042.
7	8005.	30845.	2834521.	45515.	115.	439.	87087.	679.	3007206.
8	3440.	26411.	33061.	0.	47.	376.	497.	0.	63632.
9	1681654.	16988104.	35369208.	1463655.	92514.	955115.	2090751.	41829.	58682930.

FIGURE J6

*** TRANSITION RATES MATRIX ***

THE PROBABILITY OF BEING IN A PARTICULAR FINAL STATE-LOCATION GIVEN INITIAL STATE-LOCATION
(EACH ELEMENT OF THE ACCOUNTS MATRIX DIVIDED BY ITS ROW TOTAL)

ROW NUMBER   BLOCK OF COLUMNS NUMBER 1   COLUMN NUMBERS 1 TO 8

1	0.2463198	0.0369709	0.0379767	0.0234013	0.0524515	0.0010599	0.0011490	0.0006709
2	0.0057420	0.8668747	0.0334485	0.0387722	0.0001665	0.0528626	0.0010122	0.0011213
3	0.0021472	0.0165380	0.9035792	0.0198610	0.0000623	0.0004743	0.0567395	0.0005974
4	0.0429590	0.4927819	0.4356953	0.0000000	0.0012459	0.0141330	0.0131849	0.0000000
5	0.8820357	0.0359064	0.0306022	0.0188841	0.0268329	0.0005091	0.0004636	0.0002660
6	0.0041602	0.9139268	0.0245394	0.0304323	0.0000599	0.0270909	0.0003538	0.0004368
7	0.0026619	0.0102570	0.9425763	0.0151353	0.0000382	0.0001460	0.0289594	0.0002258
8	0.0509128	0.4150585	0.5195656	0.0000000	0.0007386	0.0059090	0.0078105	0.0000000

FIGURE J7

*** ADMISSION RATES [illegible] ***

THE PROBABILITY OF [illegible] IN A PARTICULAR INITIAL STATE-LOCATION GIVEN FINAL STATE-LOCATION
(EACH ELEMENT OF ACCOUNTS MATRIX DIVIDED BY ITS COLUMN TOTAL)

ROW NUMBER BLOCK OF COLUMNS NUMBER 1 COLUMN NUMBERS 1 TO 8

ROW								
1	[illegible]	[illegible]	0.0016567	[illegible]	0.3906544	0.0017432	0.0008633	[illegible]
2	[illegible]	[illegible]	0.0157920	0.4523695	0.0300603	0.9242227	0.3080642	[illegible]
3	[illegible]	0.0327551	0.0285075	[illegible]	0.0223094	0.0172335	0.9418266	0.4756509
4	[illegible]	0.0325255	0.0128123	0.0000000	0.0151764	0.0165917	[illegible]	0.0000000
5	[illegible]	0.0002784	[illegible]	0.0016972	0.0381672	[illegible]	0.0000792	0.0008367
6	[illegible]	[illegible]	[illegible]	[illegible]	0.0008022	0.0393853	0.0002744	[illegible]
7	[illegible]	[illegible]	[illegible]	0.0410763	0.0012431	[illegible]	[illegible]	[illegible]
8	[illegible]	0.0015547	[illegible]	0.0000000	[illegible]	0.0003937	0.0002377	0.0000000

FIGURE 38

*** GROWTH RATES MATRIX(TRANSPOSED) ***

ROW NUMBER BLOCK OF COLUMNS NUMBER 1 COLUMN NUMBERS 1 TO 4

ROW				
1	0.9203818	[illegible]	0.0023778	0.0452485
2	0.0399787	0.9426784	0.0174266	0.5765363
3	0.0405407	0.0354000	0.0232527	0.4055864
4	[illegible]	[illegible]	0.0277734	0.0000000

FIGURE 39

FINAL HISTORICAL INFORMATION FROM THE REVISED ACCOUNTS MATRIX
****

(E)SURVIVING MIGRANT TOTALS

	IMMIGRATION			EMIGRATION		
	II	CI	*I	II	IO	I*
EA	170401.	48169.	218570.	117738.	36762.	154500.
SE	632023.	552546.	1184569.	654434.	647447.	1301881.
RB	618208.	488536.	1106744.	648460.	689296.	1337756.
RTW	1373205.	0.	0.	1089251.	0.	0.

(B)SURVIVING MIGRANT TOTALS

	IMMIGRATION			EMIGRATION		
	II	CI	*I	II	IO	I*
EA	13767.	3240.	17007.	8752.	2485.	11237.
SE	35570.	26411.	61981.	38365.	42150.	80515.
RB	36030.	33061.	69091.	38850.	45516.	84365.
RTW	90150.	0.	0.	62712.	0.	0.

	BIRTH		DEATH		MIGRATION+DEATH			
	AT RISK	RATE	AT RISK	RATE	(B)IMMIG	(E)IMMIG	(B)EMIG	(E)EMIG
EA	1549563.00	0.08492	1615887.75	0.05725	245.00	6340.00	163.00	4524.00
SE	16129991.00	0.08527	16803166.50	0.05662	872.00	33972.00	1178.00	38406.00
RB	33827360.00	0.08896	35305191.87	0.05922	1048.00	33451.00	1233.00	39355.00
RTW	0.00	0.00000	0.00	0.00000	1319.00	46510.00	920.00	32025.00

FIGURE J10

**** ANALYSIS OF REGIONAL POPULATION CHARACTERISTICS ****

GROUPS 1-4 INCLUDE REGIONS WITH AN INCREASE IN POPULATION
GROUPS 5-8 INCLUDE REGIONS WITH A DECREASE IN POPULATION

REGION	BIRTHS	DEATHS	IMMIGRATION	OUTMIGRATION	NAT.INCR	NET MIG	TOTAL CHANGE
+++ GROUP 1..NET IMMIGRATION EXCEEDS NATURAL LOSS +++							
+++ GROUP 2..NET IMMIGRATION EXCEEDS NATURAL INCREASE +++							
EA	131371.0	92514.0	242164.0	170424.0	39077.0	71738.0	110815.0
+++ GROUP 3..NATURAL INCREASE EXCEEDS NET IMMIGRATION +++							
+++ GROUP 4..NATURAL INCREASE EXCEEDS NET OUTMIGRATION +++							
SE	1365042.0	955110.0	1281404.0	1421980.0	409926.0	-140576.0	269350.0
RB	3007207.0	2090751.0	1210974.0	1462709.0	916456.0	-251735.0	664721.0
+++ GROUP 5..NET OUTMIGRATION EXCEEDS NATURAL INCREASE +++							
+++ GROUP 6..NET OUTMIGRATION EXCEEDS NATURAL LOSS +++							
+++ GROUP 7..NATURAL LOSS EXCEEDS NET OUTMIGRATION +++							
+++ GROUP 8..NATURAL LOSS EXCEEDS NET IMMIGRATION +++							

THE DISTRIBUTION OF POPULATION IN THE MAIN SYSTEM

REGION	OLD POP	%	NEW POP	%	% CHANGE IN POP
EA	1560533.0	2.936	1681754.0	3.112	7.748
SE	16656320.0	31.321	16988104.0	31.437	2.029
RB	34948572.0	65.743	35369208.0	65.451	1.204
	53156432.0		54039066.0		

FIGURE J11

FULLY AGGREGATE [illegible] COMPONENTS-BASED MODEL FOR 5 YEAR FORECASTING PERIOD COMMENCING 1979

MIGRATION DATA-TYPE [illegible] 5

INPUT DATA
****

	INITIAL POPULATION	BIRTH RATE	DEATH RATE
EA	1681734.	0.072947	0.057253
SE	16928104.	0.069412	0.056825
RB	35369207.	0.072238	0.050239
RTW	0.	0.000000	0.000000

(S)IMMIGRATION VECTORS

RTW

48169. 552546. 488536. 0.

(E)EMIGRATION VECTORS

RTW

30762. 647447. 689296.

EXIST AND SURVIVE TRANSITION RATES MATRIX

EA	0.0000000	0.0369769	0.0379767	0.0000000
SE	0.0057420	0.0000000	0.0334485	0.0000000
RB	0.0021472	0.0165380	0.0000000	0.0000000
RTW	0.0000000	0.0000000	0.0000000	0.0000000

EXIST AND SURVIVE MIGRATION MATRIX(DERIVED FROM INPUTS)

EA	0.	62170.	63867.	30762.
SE	97546.	0.	362227.	647447.
RB	73963.	584937.	0.	689296.
RTW	48169.	552546.	488536.	0.

(E)MIGRATION+SURVIVAL TOTALS

	IMMIGRATION			EMIGRATION		
	II	CI	+I	II	IO	I+
EA	173409.	48169.	221650.	126043.	30762.	162805.
SE	647113.	552546.	1199659.	569774.	647447.	1313221.
RB	632095.	488536.	1120631.	660030.	689294.	1350176.
RTW	1373305.	0.	0.	1289251.	0.	0.

(B)IMMIGRATION VECTORS

RTW

3260. 46411. 33661. 0.

(B)EMIGRATION VECTORS

RTW

4485. 42159. 45575. 0.

BIRTH AND SURVIVAL TRANSITION RATES MATRIX

EA	0.0000000	0.0359864	0.0366042	0.0000000
SE	0.0061602	0.0000000	0.0435194	0.0000000
RB	0.0026610	0.0102577	0.0000000	0.0000010
RTW	0.0000000	0.0000000	0.0000000	0.0000000

FIGURE J12

INFORMATION CALCULATED BY THE MODEL ON ITERATION NUMBER 5
****

****CONVERGENCE ACHIEVED ON THIS ITERATION****

BIRTH AND SURVIVAL MIGRATION MATRIX

EA	0.0	6360.5	3216.3	2425.0
SE	4753.7	0.0	26897.6	42155.0
RB	6585.4	23367.2	0.0	45215.0
RTW	1240.0	20411.0	33661.0	0.0

(M)IGRATION AND SURVIVAL TOTALS

	IMMIGRATION			EMIGRATION		
	II	CI	BI	II	IO	IA
EA	11337.	3260.	14377.	8077.	2455.	10532.
SE	29728.	26419.	56139.	31651.	44459.	73804.
RB	30014.	33661.	63675.	31951.	45315.	77466.
RTW	90150.	0.	0.	66712.	0.	0.

	BIRTH		DEATH		MIGRATION DEATH			
	AT RISK	TOTAL	AT RISK	TOTAL	(B)IMMIG	(E)IMMIG	(B)EMIG	(E)EMIG
EA	1664769.14	121429.84	1729641.75	98707.72	416.15	6427.42	155.64	4779.57
SE	16466843.06	1142664.72	17019519.32	967126.07	953.26	34576.42	1072.79	34574.34
RB	34236139.24	2473150.32	35421237.60	2199397.96	969.73	33600.15	1136.72	39770.79
RTW	0.00	0.00	0.00	0.00	1319.76	40444.71	917.70	32047.02

	(E)NON-MOVER TERMS		(B)NON-MOVER TERMS		FINAL POPULATION
	K(I,I)	K(I,N+1)	K(N+1,I)	K(N+1,N+1)	
EA	1625296.32	88873.11	107440.37	3275.03	1764979.467
SE	14735519.80	900791.93	1036936.21	30954.66	17028151.872
RB	31585488.37	1993702.23	4323542.10	71005.85	35493336.186
RTW	0.00	0.00	0.00	0.00	0.000

EXISTANCE AND DEATH MATRIX

EA	88875.1100	1774.0174	1910.5122	1067.6440
SE	4832.5497	900791.9264	17077.3068	18660.5265
RB	2205.5684	16858.8864	1993762.2338	20716.5404
RTW	1398.5267	15925.3179	14682.7716	0.0000

BIRTH AND DEATH MATRIX

EA	3275.0340	62.3886	58.4299	39.8240
SE	68.3306	30954.6565	461.1844	603.0720
RB	94.9062	302.9474	71005.8524	678.8676
RTW	40.7096	372.8622	493.1131	0.0000

FIGURE J13

AUTO-PROJECTION(Rogers G-matrix model)RUN : a base-run of the historical a.b.m. for 1965-71, followed by a run of the Rogers G-matrix forecasting model for 1971-76.

CONTROL CARDS I	A,B,C,D,E,F	LINES 0-6
DATA CARDS I	LINES 7-19	Historical data
CONTROL CARDS II	G,H	LINES 20+21

```
 0    4  4 -1  1 10  0
 1    1
 2 1971 5
 3    3200
 4    EA  SE  RB RTW
 5        A FOUR REGION SEMI-CLOSED SYSTEM EXAMPLE USING MULTIPOP2
 6 1966-71 ACCOUNTS FOLLOWED BY 1971-76 FORECAST WITH ROGERS MODEL
 7 1560535 16650328 34948572 0
 8 0 51397 53941 34315
 9 108034 0 566655 678122
10 83321 565593 0 716485
11 51060 516188 466291 0
12 131591 1385042 3007207 0
13 92514 955116 2090751 0
14 0 4202 3659 2331
15 6519 0 33211 44326
16 8887 30178 0 46971
17 3425 24608 31473 0
18 1570938 16698754 34704488 1121280 0 0 0 63631
19 1681754 16988102 35369209 1463655 0 0 0 41830
20    0
21 2
22 000
```

TABLE 14 : INPUT FILE FOURREGION2

ROGERS G-MATRIX FORECASTING MODEL FOR PERIOD 1971-1976
****

SEMI-CLOSED SYSTEM--IMMIGRATION VECTOR

****
* INPUT DATA *
****

1. OLD G-MATRIX USED

2. SURVIVING IMMIGRANTS

EA	51409.
SE	578957.
RB	521597.

****
* RESULTS *
****

POPULATION TOTALS

REGION	INITIAL	PROJECTED	% CHANGE
EA	1681754.	1787111.	6.26
SE	16988104.	17276883.	1.70
RB	35369208.	36038863.	1.89
RTW	0.	0.	0.00

THE DISTRIBUTION OF POPULATION IN THE MAIN SYSTEM(%)

REGION	INITIAL	PROJECTED
EA	3.1121	3.2432
SE	31.4367	31.3539
RB	65.4512	65.4029
	***********	***********
TOTAL	54039066.	55102857.

FIGURE K : OUTPUT FROM A TYPICAL AUTO-PROJECTION(ROGERS G-MATRIX) RUN OF MULTIPOP2, USING DATA FILE FOURREGION2 - RESULTS FOR A FORECASTING PERIOD.

References

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

Hamilton, C.H. (1967) The vital statistics method of estimating net migration by age cohorts. Demography, 4 (2), pp. 464-78.

Illingworth, D.R. (1976) Testing some new concepts in population accounting and related fields. Ph.D. dissertation, School of Geography, University of Leeds.

Jenkins, J.C. (1976) Problems of assembling international migration data from censusesand other sources for use in migration models, population accounts and forecasts: a preliminary exploration. Working Paper 161, School of Geography, University of Leeds.

Macgill, S.M. (1975) Balancing factor methods in urban and regional analysis. Working Paper 124, School of Geography, University of Leeds.

Rees, P.H. (1974) A family of demographic accounts. Working Paper 68, School of Geography, University of Leeds.

Rees, P.H. (1976) Modelling the regional system: the population component. Working Paper 148, School of Geography, University of Leeds.

Rees, P.H. (1977a) The measurement of migration from census data and other sources. Environment and Planning A, 9 (3), pp. 247-272.

Rees, P.H. (1977b) The future population of East Anglia and its constituent counties (Cambridge, Norfolk and Suffolk). A report prepared for the East Anglian Economic Planning Council by the University of Leeds Industrial Services Limited, under Department of the Environment contract number DGR/461/23, "East Anglia: research into population".

Rees, P.H. and Jenkins, J.C. (1977) Let a hundred flowers bloom and a hundred schools contend. Working Paper in prepatation.

Rees, P.H. and Wilson, A.G. (1974) A computer programme for constructing spatial demographic accounts: user's manual. Working Paper 60, School of Geography, University of Leeds.

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

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

Webb, J.W. (1963) Natural and migrational components of population changes in England and Wales, 1921-1931, Economic Geography, 39(2), pp. 130-148.

Appendix A: Notes for the Users of the Leeds 1906A Machine-Listing Files and Running Jobs.

Users with access to the University of Leeds 1906A machine will be able to produce their own listings of the three programmes, and run the programmes under the GEORGE operating system. The following instructions are necessary to complete these operations.

(a) LISTING PROGRAMMES A numbered listing may be produced either online or offline as follows:-

Online

```
LN‸jobname,:username
LF‸:GEOLIB.MULTIPOP2,*LP,NU
LT
```

Offline

```
JOB‸jobname,:username
LF‸:GEOLIB.MULTIPOP2,*LP,NU
EJ
```

(b) CREATING DATA FILES Users may run the programmes with the input information being read from previously created data files. The instructions for creating such a file are as follows:-

```
INPUT‸:username,datafilename
Input data (eg: as in Table 9)
****
```

(c) RUNNING THE PROGRAMMES The programmes may be run either online or offline. If they are to be run online the input information must be read from a file which has already been created (as instructed in part b):-

```
LN‸jobname,:username
RJ‸jobname,PROG,JD(JT‸30,MZ‸150K),PARAM-
(BIN‸:GEOLIB.MULTIPOP2BIN,DATA datafilename)
LT
```

Jobs may be run offline with the inputs being read in directly after the job instructions:-

```
JOB‸jobname,: username, JD (JT‸30,MZ‸150K)
PROG‸BIN‸: GEOLIB.MULTIPOP2BIN,DATA,TL 25
input data (eg: as in Table 9)
****
```

or again with the inputs being read from a previously created data file:-

```
JOB‸jobname,:username,JD(JT‸30,MZ‸150K)
PROG‸BIN‸:GEOLIB.MULTIPOP2BIN,DATA‸datafilename, TL‸25
EJ
```

In the above instructions MULTIPOP2BIN is the binary version of the MULTIPOP2 programme, and we advise its use when running jobs as this considerably cuts down the time required by each job. As well as reading data from a file, the results of any programme run may be saved in a file and later listed. This facility is particularly useful if users require several copies of the results. If users do wish to save the output of a run in a file the instructions above should be modified as follows:-

<u>DATA‸datafilename</u> should be replaced by

<u>FILE*CRO=datafilename,FILE*LPO=outputfilename</u>

In the instructions in parts (a), (b) and (c) we referred to MULTIPOP2, but if users prefer to use either the MULTIPOP1 or MULTIPOP3 programmes (or their binary versions) they may do so simply by inserting the new programme name in the appropriate position.

(d) <u>NOTES ON JOB-TIMES FOR MULTIPOP2</u>. We have advised the use of the binary versions of any programme as this considerably reduces the amount of the user's computing budget consumed by any one job. For example, the first four-region example in section 9 required 41 seconds to compile and run, but using the binary version of MULTIPOP2 (MULTIPOP2BIN) reduced the job-time to 5 seconds. In addition to saving computer time also ensures results are returned more quickly. We have run a series of examples using MULTIPOP2BIN to provide users with some idea of the time taken by various sized jobs. Hopefully these examples will allow users to specify job times in the job command statements (i.e. the JT parameter) more accurately. In the instructions we have suggested a job time of 30 seconds and generally this will be sufficient for most purposes. All the job times to follow relate to examples in which only the final results are output. Examples were run for two regional systems of four and fifteen zones, and the times required were as follows in seconds of computer time:-

	Four region system	Fifteen region System
Single-period base-run	4	7
Historical a.b.m.		
+ 1 a.b.m. forecast	4	11
+ 3 "	6	19
+ 6 "	9	30
+ 10 "	Not run	45
Historical a.b.m.		
+ 6 Rogers-model forecasts	5	8

The population constraints routine, which may be used in the historical modelling, added a further second to the times of the four-region system examples.

Appendix B: Infant Deaths" Estimation

In the semi-aggregate models described in sections 4 and 5 of the paper, which may be run in the MULTIPOP1 and MULTIPOP3 programmes, two deaths sub-totals are required as inputs rather than the single deaths total of the fully-aggregate models. In other words the single total must be subdivided into those who were alive at the start of the modelling period and those who were not (ie: who were born later).

The DETHINF programme has been written to do this, since the information is generally not directly available in this form. It allocates deaths by age group to the birth and existence cohorts, to produce $K^{\beta(*)\delta(i)}$ and $K^{\varepsilon(*)\delta(i)}$ terms respectively, using a set of geometrically - defined probabilities similar to those used, for example, by Hamilton (1967). These probabilities indicate the likelihood that persons aged x days or years at their deaths have been born during the modelling period. Clearly we need only concern ourselves, however, in the present models with those persons dying under the age of y years in any period, where y is the length of the period.

For the task at hand it is essential that the data on deaths under one year of age, at least for the first year in the modelling period if that period is greater than one year in length, is as disaggregated as possible in terms of the age-group categories used in the estimations. This is because deaths in the first year of life are very unevenly distributed within the age group, the majority being concentrated in the "less-than-one-day-old" group. Virtually all of these deaths in any one year will be of infants born in the same year. The more age groups that are used, therefore, up to a maximum of about 25 beyond which results are not significantly changed, the less underestimated is the number of deaths allocated to the birth cohort, particularly if deaths under one week are split into six or seven categories. In all other age groups, other than the under-one-year group, as assumption that deaths are evently distributed within the group seems reasonable, and has been adopted.

The DETHINF programme may be used to calculate the deaths sub-totals for a one or a five-year period. The instructions for preparing the inputs to the programme are set out in Table B(1). Table B(2a) gives a typical input listing for a one-year period example, with data on the deaths of Danish boys in years 1971 and 1972, and Table B(2b) gives the results output for this example. A full listing of the DETHINF programme then follows.

For a one-year period the only data that needs to be input are the number of deaths of those under one in the year in question by the selected age groups, and the total number of all deaths in that year. The deaths total is used to calculate the number of deaths in the existence cohort through eliminating the infant deaths as estimated in the programme. In a five-year period case an additional set of figures needs to be input. For each of the five years data on the death of under fives is required. This may be input in two forms:-

(a) deaths in five single-year age groups; or
(b) deaths in the under one, and 1-4 age groups.

The choice is indicated through the IY parameter in Input 2. The deaths data in Input 6 must then be read in as a five by five matrix, if the first form is selected, with deaths under one as the first row and the first year of the period as the first column, or as a two by five matrix, if the second form is chosen.

Inputs 4 to 7 may be repeated in sequence for as many examples as there are to be run, as specified in parameter IJC in Input 2 (as in the example in Table B(2a) where two sets of estimations are carried out).

INPUT 1: TITLE CARD. Eighty columns - any alphanumeric characters.
INPUT 2: PARAMETER CARD (Format 4I2)

1-2	IY	One-year period - any two digits may be inserted in these columns. Five-year period - a value of 5 must be used if the deaths data to be read in refers to single-year age groups; or a value of 2 if the data refers to 0-1 and 1-4 year age groups.
3-4	NA	The length of the period in years: 1 or 5 only.
5-6	L	Number of age groups for which data on deaths under one is to be input - maximum 24.
7-8	IJC	Number of examples to be run.

INPUT 3: THE AGE GROUPS FOR DEATHS UNDER ONE (Format 24F3.0)
(a) the upper boundaries of each age group in days
(b) the lower boundaries of each age group in days
Each should start with the youngest group, and progress in sequence to the oldest.

INPUT 4: THE YEAR TO WHICH THE DATA REFERS - FOR ONE-YEAR PERIOD EXAMPLES ONLY (Format I4).

INPUT 5: DATA BY AGE GROUPS FOR DEATHS UNDER ONE (Format 12F4.0). Twelve values per line in the order specified in Input 3.

INPUT 6: DATA BY AGE GROUPS FOR ANNUAL DEATHS UNDER FIVE (Format 10F6.0). FOR FIVE-YEAR PERIOD EXAMPLES ONLY. Either a 5 x 5 data matrix if the IY parameter has been set to 5, or a 2 x 5 data matrix if IY equals 2.

INPUT 7: TOTAL DEATHS OVER THE PERIOD (Format F9.0)

TABLE B(1): INPUT INSTRUCTIONS FOR THE DETHINF PROGRAMME.

```
0 DEATHS UNDER ONE YEAR FOR DENMARK 1971+72..MALES..21 GROUPS
1 99 121 2
2    0   1   2   3   4   5   6 13 20 27 59 90120151181211242273303333364
3    0   1   2   3   4   5   6   7 14 21 28 60 91121152182212243274304335
4 1971
5  210   90   50   29   20   10   12   30   21   11   33   11
6    19   13   15    6    6    5    4    4    3
7      26837
8 1972
9   204   95   45   24   15   10    9   25   15    6   25   16
10   18   16    7    4    7    4    5    7    4
11     27482
12 ****
```

TABLE B(2a) : INPUT LISTING FOR THE DETHINF PROGRAMME

```
DEATHS UNDER ONE YEAR FOR DENMARK 1971+72..MALES..21 GROUPS

ONE YEAR PERIOD

    CALCULATIONS FOR YEAR  1971

DEATHS ALLOCATED TO THE BIRTH COHORT IN THE FIRST YEAR          553.9

DEATHS ALLOCATED TO THE BIRTH COHORT FOR  THE PERIOD          553.9

DEATHS ALLOCATED TO THE EXISTENCE COHORT FOR THE PERIOD          26253.1

    CALCULATIONS FOR YEAR  1972

DEATHS ALLOCATED TO THE BIRTH COHORT IN THE FIRST YEAR          513.5

DEATHS ALLOCATED TO THE BIRTH COHORT FOR  THE PERIOD          513.5

DEATHS ALLOCATED TO THE EXISTENCE COHORT FOR THE PERIOD          26968.5
```

TABLE B(2b) : OUTPUT PRODUCED BY THE DETHINF PROGRAMME FROM THE INPUT EXAMPLE IN TABLE B(2a)

FULL LISTING OF THE DETHINF PROGRAMME

```
         LIST
         PROGRAM(YDYD)
         TRACE 2
         INPUT 5=CR0
         OUTPUT 6=LP0
         END
         MASTER YDED
C A PROGRAM TO ALLOCATE DEATHS TO THE BIRTH COHORT FOR A ONE OR
C FIVE YEAR PERIOD USING ONE OR FIVE YEAR DATA,AND DISAGGREGATED
C INFORMATION ON DEATHS UNDER ONE YEAR OF AGE IN THE FIRST YEAR
C OF THE PERIOD .
         DIMENSION TITLE(10),E(25),B(25),PROB(25),YD(25),YDEATH(10,
        110)
         READ(5,1)TITLE
       1 FORMAT(10A8)
         WRITE(6,2)TITLE
       9 FORMAT(10F6.0)
       2 FORMAT(1H1,10A8)
         READ(5,3)IY,NA,L,IJC
       3 FORMAT(4I2)
         IF(NA.EQ.1)WRITE(6,111)
         IF(NA.EQ.5)WRITE(6,112)
     111 FORMAT(/1H ,'ONE YEAR PERIOD')
     112 FORMAT(/1H ,'FIVE YEAR PERIOD')
         READ(5,4)(E(I),I=1,L)
         READ(5,4)(B(I),I=1,L)
       4 FORMAT(24F3.0)
         DO 25 I=1,L
      25 PROB(I)=1-(((2*B(I))+(1+E(I)-B(I)))/730)
         IJCC=0
      16 IJCC=IJCC+1
         DETHIN=0.0
         IF(NA.EQ.1)READ(5,17)IYEAR
      17 FORMAT(I4)
         IF(NA.EQ.1)WRITE(6,18)IYEAR
      18 FORMAT(/1H0,'     CALCULATIONS FOR YEAR   ',I4/)
         READ(5,5)(YD(I),I=1,L)
       5 FORMAT(12F4.0)
C CALCULATION OF DEATHS IN THE BIRTH COHORT IN YEAR ONE
         DO 26 I=1,L
         YD(I)=YD(I)*PROB(I)
      26 DETHIN=DETHIN+YD(I)
         WRITE(6,6)DETHIN
       6 FORMAT(/1H ,'DEATHS ALLOCATED TO THE BIRTH COHORT IN THE',
        1' FIRST YEAR ',F11.1)
         IF(NA.EQ.1)GO TO 10
       8 DO 7 J=1,IY
       7 READ(5,9)(YDEATH(J,I),I=1,NA)
         IF(IY.EQ.2)GO TO 100
```

(continued)

```
C ESTIMATION OF DEATHS IN BIRTH COHORT USING ONE YEAR AGE GROUPS
        NAS=NA+1
   30   I=2
        J=2
        DETHIN=DETHIN+0.5*YDEATH(I,J)
        I=I+1
        J=J+1
        IF(I.LT.NAS)GO TO 30
        NAU=NA-1
        IR=1
        IC=IR+1
        DO 67 II=IR,NAU
        DO 67 ID=IC,NA
        IF(II.LT.ID)GO TO 67
        YDEATH(II,ID)=0.0
   67   DETHIN=DETHIN+YDEATH(II,ID)
        GO TO 10
C CALCULATION OF DEATHS IN THE BIRTH COHORT USING AGE GROUP
C DATA 0-1 AND 1-4
  100   DO 60 J=2,NA
   60   DETHIN=DETHIN+YDEATH(1,J)
        JZ=1
        DO 61 J=2,NA
        YDEATH(2,J)=YDEATH(2,J)*(0.125*JZ)
        JZ=JZ+2
   61   DETHIN=DETHIN+YDEATH(2,J)
C OUTPUT OF RESULTS FOR THE WHOLE PERIOD
   10   WRITE(6,40)DETHIN
   40   FORMAT(/1H ,'DEATHS ALLOCATED TO THE BIRTH COHORT FOR ',
       1' THE PERIOD ',F11.1)
        READ(5,41)DEATH
   41   FORMAT(F9.0)
        EDETH=DEATH-DETHIN
        WRITE(6,42)EDETH
   42   FORMAT(1H0,'DEATHS ALLOCATED TO THE EXISTENCE COHORT FOR',
       1' THE PERIOD ',F12.1)
        IF(IJCC.LT.IJC)GO TO 16
        STOP
        END
        FINISH
****
```

APPENDIX C: A NOTE ON THE MULTIPOP1 AND MULTIPOP3 PROGRAMMES.

The range of modelling options available in the MULTIPOP1 and MULTIPOP3 programmes was spelt out briefly in Part I of the paper and summarised in Table 1. Details of the models themselves were given in the various sections of Part II. Here the close relations between the different models were emphasized, and this similarity extends also to the three computer rrogrammes that have been developed. We feel sure that readers who have digested Part III of the paper on the MULTIPOP2 programme will easily be able to master the details of the two remaining programmes without our having to produce lengthy discussions of their form, input requirements and outputs. Structurally they are very similar in terms of their subroutine format and functions. The inputs vary slightly in that extra parameters and data are sometimes required, but the input listings of MULTIPOP1 closely resemble those of Tables 8 and 11, and those of MULTIPOP3 the whole suite of inputs described in section 8. Despite all the similarities, we hope we have extended the range of operational accounts-based models in some useful directions.

May we repeat that readers who would like to use the MULTIPOP1 and MULTIPOP3 programmes are welcome to write to the authors at the School of Geography for listings and/or machine readable copies of the programmes in which they are interested. We shall be pleased to provide these (and to supply additional user notes on the two programmes not fully described in this paper) at reproduction and mailing cost.

ADDENDUM: Choice of equations for calculating the population at-risk of giving birth.

A choice of alternative forms of at-risk population equations for births is now available in the MULTIPOP2 programme. In addition to the equations for calculating "multi-regional" populations at-risk of giving birth that are outlined in the main text (equation 3.2 gives the initial estimate, and 3.15 the subsequent estimates), a second means of arriving at the at-risk populations has been included. This is the "average" population at-risk of giving birth (that is, the average population of the appropriate region in the modelling period):-

$$\hat{K}^{Bi*} = 0.5\,(K^{\varepsilon(i)*(*)} + K^{*(*)\sigma(i)})$$

The initial population of region i, $K^{\varepsilon(i)*(*)}$, is used as the initial estimate of this term for region i.

The choice as to which form of at-risk population is required is made through an extra parameter that has been added to card A (see page 53) which is:-

Columns	Parameter	
19-21	NBIR	A value of 0 indicates that the "multi-regional" form is required; and a value of 1 the "average" form.

If no value is specified for this parameter the multi-regional form will be used by default.

We recommend that the "average" form be used only in runs in which forecasting is to take place using externally-derived crude birth rates (the number of births per thousand population) as inputs to the programme. In such a case the "multi-regional" population at risk of giving birth would be incompatible with the rate with which it is to be used in the forecasts. This point is discussed in a more general context and at greater length in Rees and Jenkins (1977).