A PROJECTION MODEL FOR SMALL AREA POPULATIONS: APPLICATION TO WEST YORKSHIRE DISTRICTS

Philip Rees

WORKING PAPER 91/10

SCHOOL OF GEOGRAPHY • UNIVERSITY OF LEEDS

WORKING PAPER 91/10

A PROJECTION MODEL FOR SMALL AREA POPULATIONS: APPLICATION TO WEST YORKSHIRE DISTRICTS

PHILIP REES

DECEMBER 1991

GMAP LTD
University of Leeds
3F Springfield House
Hyde Terrace
LEEDS
LS2 9TN

School of Geography
University of Leeds
LEEDS
LS2 9JT

CONTENTS

ABSTRACT

The paper describes the main features of a model designed to project the populations of small areas (wards) in England and Wales, with illustrations of the methods based on an implementation for 3 districts in West Yorkshire. The organisation and rationale of the projection model are discussed in section 2 of the paper, while sections 3 and 4 review the methods used to estimate inputs to the model. The third section treats "vector" inputs, namely mortality, fertility and emigration rates and probabilities, and immigration flows for the wards. The fourth section describes "matrix" inputs involving internal migration between the district and the rest of the country or between wards within the district. Section 5 outlines the nature of the software that implements the projection model, while the final section describes further stages in its development consequent on the publication of the 1991 Census Small Area and Local Base Statistics.

ACKNOWLEDGMENTS

This paper has been prepared under a contract betwen GMAP Ltd, University of Leeds and the metropolitan districts of Bradford, Calderdale and Kirklees in West Yorkshire to develop a projection model for small area populations in each authority. The financial support of the authorities is gratefully acknowledged.

The author is also grateful for the support of the School of Geography in allowing him the time to work on the application of his ideas in this project.

Steve Simpson earns my thanks for his valuable comments on the algorithm for estimating ward populations originally developed for Calderdale in connection with student GIS projects and here generalised to work with any English or Welsh district.

LIST OF TABLES

LIST OF FIGURES

O that a man might know
The end of this day's business ere it come!
(William Shakespeare, Julius Caesar (1599-1600), 5.1. 122)

1. OVERVIEW

1.1 Introduction

There is an increasing interest, expressed by both public authorities and private firms, in being able to forecast the demand for goods and services in flexibly defined but relatively small scale areas over the short term future, the next ten years. One of the ingredients in such a forecast is good knowledge of the likely sizes and demographic structures of populations living in small areas. Of course, for proper forecasting of the demand for public or private goods and services, it is necessary to know something of the socioeconomic make-up of the population and of the facilities currently supplying the goods or services being studied. However, any such an analysis must build from a reliable projection of the client or customer population. This paper outlines the principal features of a population projection model currently being constructed for three local authorities in West Yorkshire: the metropolitan districts of Bradford, Calderdale amd Kirklees. The projection model, however, has been specified in a general way and could easily be adapted for use with any English or Welsh districts.

The rest of this introductory section to the paper reviews the aims of the model, the characteristics of the system of interest being modelled and the broad methods used. The second section of the paper gives an account of the overall structure of the projection model. Sections 3 and 4 then describe the techniques employed to estimate, for small areas, the component inputs to the projection model. These are the base populations (section 3.1), mortality rates (section 3.2), fertility rates (section 3.3), and external migration rates and flows (sections 3.4 and 3.5) and internal migration flows, rates and associated variables are discussed in section 4. In the fifth section details are provided of the implementing software, the ways in which model outputs can be viewed (section 5.2), and the methods for developing scenarios of future component inputs (section 5.3). The final section describes further stages in projection model development consequent on the publication of new data in the 1991 Census.

1.2 Aims of the projection model

The aims of the model are to provide a tool for users to carry out their own projections of the populations of small areas within districts. The user decides on the assumptions that will be used in the projection. The model can be used, therefore, for rapid revision of projections when new information invalidates assumptions. It can also be used to develop a range of plausible futures or to explore the system's sensitivity to changes in various input assumptions. The flexibility provided by the model contrasts strongly with the rigidity of official forecasts (eg. OPCS 1991), which provide only one projection; or of commercial derivatives based on official forecasts (eg CACI's projections described in Whitehead and Dugmore 1990).

1.3 The system of interest

The system of interest can be described in terms of the entities involved, the areas in which they are counted, the characteristics of the entities and the time points and periods involved in estimation and projection.

1.3.1 The entities

The model principally uses individuals as its entities and, when counted for particular areas or categories, populations. Households are generated from knowledge of the age, gender and marital structure of populations but this part of the model is included in a second phase described elsewhere. However, population per household statistics are used in the projection model to convert new occupied housing units into counts of in-migrating individuals, and to convert housing units that are demolished into counts of out-migrating individuals.

1.3.2 The areas

The small areas, the populations of which are projected, are the wards within districts. The reason wards were selected is that statistics on electors, births, deaths and, more occasionally, migrants are published by the Office of Population Censuses and Surveys. In the second phase of the model, a method is developed for producing from ward populations estimates and projections, estimates for other, user defined areas. This will involve disaggregating ward figures to enumeration district level and reaggregating to user defined areas.

The model also involves areas external to the district of interest. These are other districts with which the target district interacts and the rest of the country, defined as the rest of the United Kingdom. Figure 7 provides lists of zones for the districts.

1.3.3 Ages and times

In order to meet local authority requirements for projected populations flexibly aggregated to any age combination and for annual projections, the projection model was designed *ab initio* using annual time periods and single years of age. It is not possible to represent all population flows in the model at this scale because of the small numbers involved. Although the model handles inter-ward and inter-district flows, age and gender disaggregation is only applied at the last stage to total outflows and inflows, for example.

Annual time intervals are of two kinds. The first are calendar years which are used mainly in the estimation of mortality, fertility and external migration rates in the intercensal period (1981 to 1990). The second are intervals from mid-year in one calendar year to mid-year in the next. This is the interval used in projection to move from one population stock to the next. Population estimates in intercensal period are made at mid-year, and the internal migration estimates refer to the mid-year intervals 1980/81 to 1989/90.

1.3.4 Estimation and projection

The term "estimation" is used to describe the process of calculating the most probable figure for a population stock or flow for a time point or period in the past. The term "projection" is used to describe the process of computing the most probable figure, based on particular assumptions, for a population stock or flow for a time point or period in the future.

Estimates are prepared for most component variables for either calendar year or mid-year intervals between 1981 and 1990, so that trends can be carefully examined before deciding on projection assumptions. However, in many cases, the numbers for the small areas involved are small and erratic, and often five year averages are therefore computed for 1981-85 and 1986-90 calendar years or 1981-86 and 1986-91 mid-year intervals.

1.4 The methods used

The projection model here embodies a large range of different techniques and so is best described as hybrid. The basic principles of multistate and multiregional population models (Rees 1989) are incorporated in a consistent way. The projection model is based on the transition concept, includes both internal and external migration and so is properly closed and includes the method of cohort survival involving period-cohort age-time plan (see Rees and Woods 1986 and Rees 1990, Chapters II, III and VI for detailed exposition of these concepts). The model marries a multiregional cohort survival model operating at the inter-district level with a set of spatial interaction models (Wilson 1974) being used at the intra-district scale. These models link migration to housing stock changes and so make possible the assessment of the population impacts of housing developments.

A variety of techniques are used to fill the information gaps in the intercensal period. Most, however, involve the combination of age-aggregated small area data with more detailed age-disaggregated national data, with resort to model schedules of age-specific rates or probabilities in the absence of adequate national data (Rees 1990, Chapters IV and V, gives a comprehensive account of estimation methods).

2. THE PROJECTION MODEL

In this section an account is given of the structure of the projection model and the procedures involved. The account is provided in verbal terms so as to be accessible to the widest audience.

Figure 1 sets out the overall structure of the projection model. It consists broadly of a loop through which the computations proceed once per time interval. At the end of each time interval the projected final population is reaccessed as the start population for the next time interval, with care being taken to move populations to the next period-cohort.

Data for the projection came from two sources: (1) The benchmark data for population components for the period since the previous census and (2) assumptions about the future development of population components, which are here called scenarios. The user of the projection model selects benchmark dataset to be used. These will refer either, on the one hand, to each calendar year or each mid-year interval or, on the other, to data averaged over periods of five years in length. Although normally data for the latest year would be selected, there may be several good reasons for choosing an alternative year or a longer period. The estimate for the latest year may embody only partial information or be subject to transitory influences (eg an influenza epidemic). The user may feel, for example, that the migration patterns of the 1981-86 period better reflect the prospects in 1991-96 than to the 1986-91 experience, given the dependence of migration behaviour on conditions in the economic cycle.

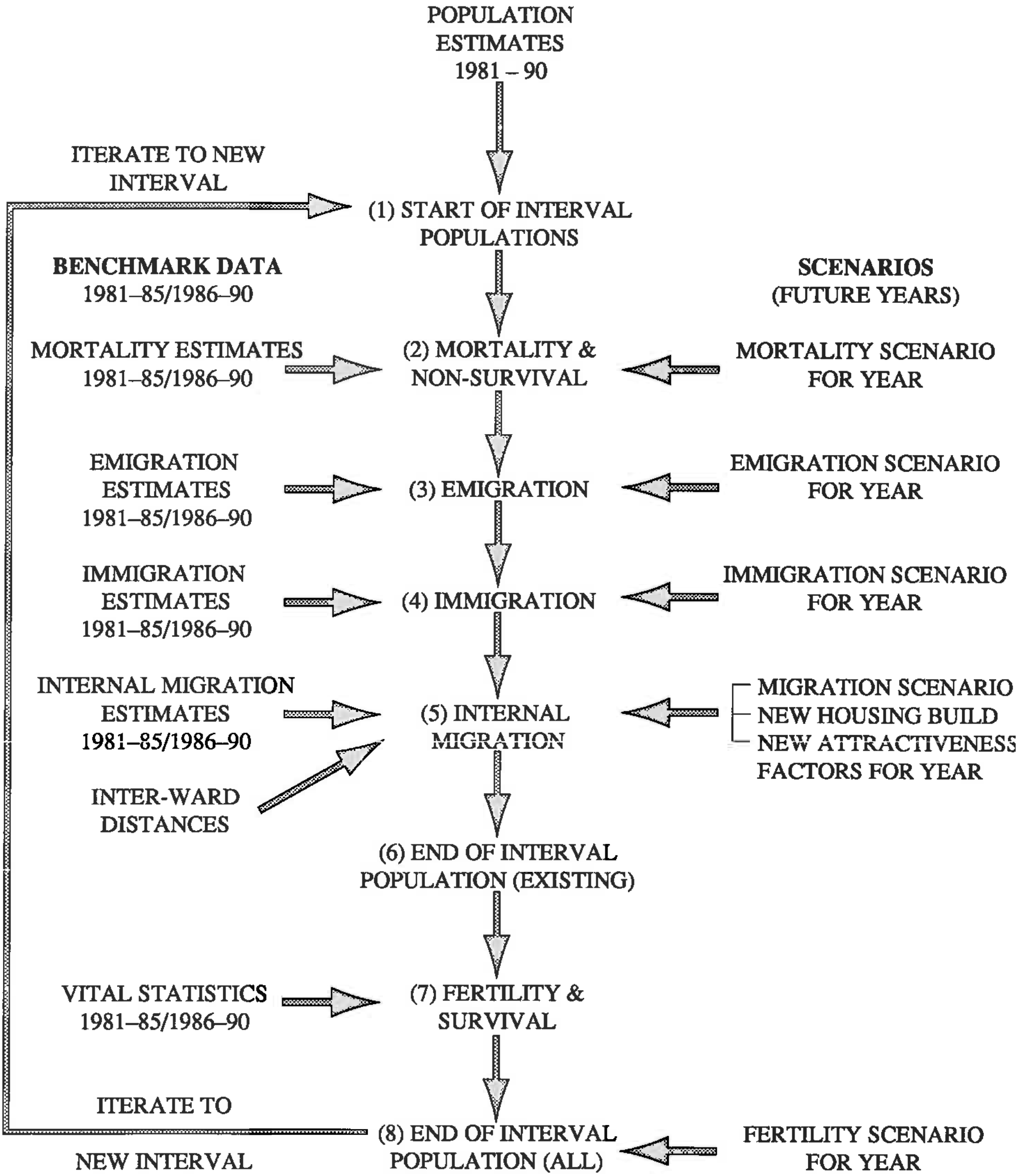

FIGURE 1. THE STRUCTURE OF THE PROJECTION MODEL

The projection model uses the following accounting equation for each period-cohort and gender

Start of interval population

- Non-survivors
- Surviving emigrants
+ Surviving immigrants
- Surviving internal out-migrants
+ Surviving internal in-migrants
= End of interval population

Each of terms refers to the projected numbers of the component population (1)

After start of interval populations for all ages and both genders have been processed, the number of projected births can be computed.

First, the populations at risk are computed by quinquennial ages

Populations at risk = the average of start of interval and end of interval female populations summed to form five year ages (2)

Second, births are projected

Births = the sum over all maternal ages of fertility rates multiplied by the population at risk (3)

The births are then entered into equation (1) as start of interval populations of "infants" in order to compute the end of interval population in the first age.

2.1 Start of interval populations

The first step in projection is the input of start of interval populations or the transfer of the previous interval's end of period populations. Input populations for each zone in the model (35 for Bradford, 23 for Calderdale and 29 for Kirklees) are broken down by gender (male, female) and by single ages 0 to 89 with 90+ as the final age (91 ages in all).

2.2 Mortality and non-survival

Mortality rates averages for period-ages are estimated for two five year periods and converted to non-survival probabilities for period-cohorts (see section 3.2 for details). Table 1 shows the period-ages and period-cohorts involved in the model. Much of the data input to the estimation procedures is organised in period-age form, but before entry to the projection model it is re-estimated in period-cohort form.

The start of interval populations in each age and gender group are multiplied by the appropriate non-survival probability to yield the projected number of non-survivors, which is subtracted from the start population.

2.3 Emigration

Emigration rate averages for period-ages are estimated for two five year periods and converted to emigration and survival probabilities (to avoid double counting of deaths) for period-cohorts (see section 3.4 for details). The start of interval populations in each age and gender group are multiplied by the emigration and survival probabilities to yield the projected number of surviving emigrants, which is subtracted from the start population.

TABLE 1. Period-ages and period-cohorts

Period-ages	Period-cohort Start age	Period-cohort Finish age
0	Birth	0
1	0	1
2	1	2
89	88	89
90+	89	90
	90+	91+

Note:

1. There will always be one more period-cohort than period-age

2.4 Immigration

Immigration flow averages for period-ages are estimated for two five year periods and converted to number of surviving immigrants for period-cohorts (see section 3.5 for details). The number of surviving immigrants is added to the start populations. Note that immigrants are treated as flows not rates because immigration is controlled by legislative and administrative rules rather than a product of ward or district populations.

2.5 Internal migration

Internal migration is handled separately for inter-district flows and for intra-district.

2.5.1 Inter-district migration

Inter-district migration probabilities are applied to district starting population to yield the number of surviving inter-district migrants. These are shared out by wards using proportions derived from the 1981 Census Special Migration Statistics for both out-migration and in-migration to yield the numbers of surviving out-migrants from wards to other districts and the numbers of surviving in-migrants from other districts.

2.5.2 Intra-district migration

Three spatial interaction models are used to project, respectively, turnover migration between wards (within the existing housing stock), migration between wards to new housing and migration between wards out of housing which is demolished. The projected migration is measured as surviving migrants in each case.

2.5.3 Total surviving migrants from and to wards

The total number of surviving out-migrants from each ward and the total number of surviving in-migrants to each ward are computed as sums.

2.5.4 Age and gender disaggregation

The out- and in-migration sums are disaggregated by age and gender through the application of model schedule migration rates adjusted to yield the sums previously projected (see section 4.4 for details). The number of surviving internal out-migrants for each period-cohort and gender is subtracted from the corresponding start of interval population while the number of surviving in-migrants for each period-cohort and gender is added to the starting population.

2.6 End of interval populations

Surviving populations in each ward at the end of the projection interval have by this stage been computed for final ages 1 to 91+.

2.7 Fertility and survival

The computation of new additions to the population through birth are most conveniently carried out after all the other period-cohorts have been processed. The method was described earlier in equations (2) and (3). Note that the new babies must be subjected to the same mortality, emigration, immigration and internal migration processes as the other period-cohorts in order to yield the end of interval populations aged 0. However the probabilities used must reflect that babies born during a projection interval are exposed, on average, for only half the interval.

2.8 Transfer of populations

At this stage all final populations for the wards have been computed. They become the starting populations for the next projection interval. Note that, before becoming the starting populations for the next projection interval, the populations in the two final ages, 90 and 91+ are summed to form a 90+ age.

The techniques and assumptions made in estimating the benchmark data are now described, beginning with the mid-year populations of wards.

3. THE ESTIMATION OF MODEL INPUTS (1)

3.1 The estimation of base populations

Although the last detailed information on the populations of wards was provided by the 1981 Census, a method has been developed to estimate ward populations by single ages and gender for any intercensal year. Although, in principle, the projection model could be used to produce ward estimates between 1981 and 1990, the difficulty is that ward population estimates are needed in order to estimate the component rates and probabilities input to the projection. It is not clear, from the published description (Whitehead and Dugmore 1990, pp 73-74) how this difficulty is overcome in the CACI model, unless cohort survival is carried out using national rates only.

Figure 2 sets out the structure of the model used to estimate ward populations. The ward populations from the 1981 Census are aged forward one year at a time (cohort survival with 100% survival), and ward births are added a year at a time to form a first estimate of the age 0 populations. In each year, the matrix of populations (wards x age/gender groups) is adjusted, using iterative proportional fitting (IPF), to fit ward totals of the electoral age (18+) age population and the population below electoral age (<18), derived from the electoral register, and to fit district totals by single age and gender derived from OPCS's mid-year estimate.

To illustrate the IPF process, consider Table 2. The interior of the table contains an initial estimate of the population across wards and age gender groups. The table marginals contain the information on total electors per ward (row totals) and on the total in each age/gender group for the district (column totals). A quick check of the numbers shows that they have to be raised, redistributed from north ward to south ward, and from the younger age group to the older. This is accomplished using IPF (see Rees 1991 for a more precise definition of the algorithm and Fenburg 1970 for the original derivation).

Although a small amount of underenumeration may take place in the course of electoral registration, this is compensated for by adjusting ward electoral age and below electoral age populations to the more reliable district total. However, there is an assumption that underregistration is uniform across wards. The influence of mortality on older age groups is allowed for by adjusting to district populations which have been cohort-survived. However, there is an assumption that mortality acts uniformly across wards. If these assumptions prove untenable, the model can be revised to incorporate adjustment for differential underenumeration and for differential cohort survival (at the expense of added complexity).

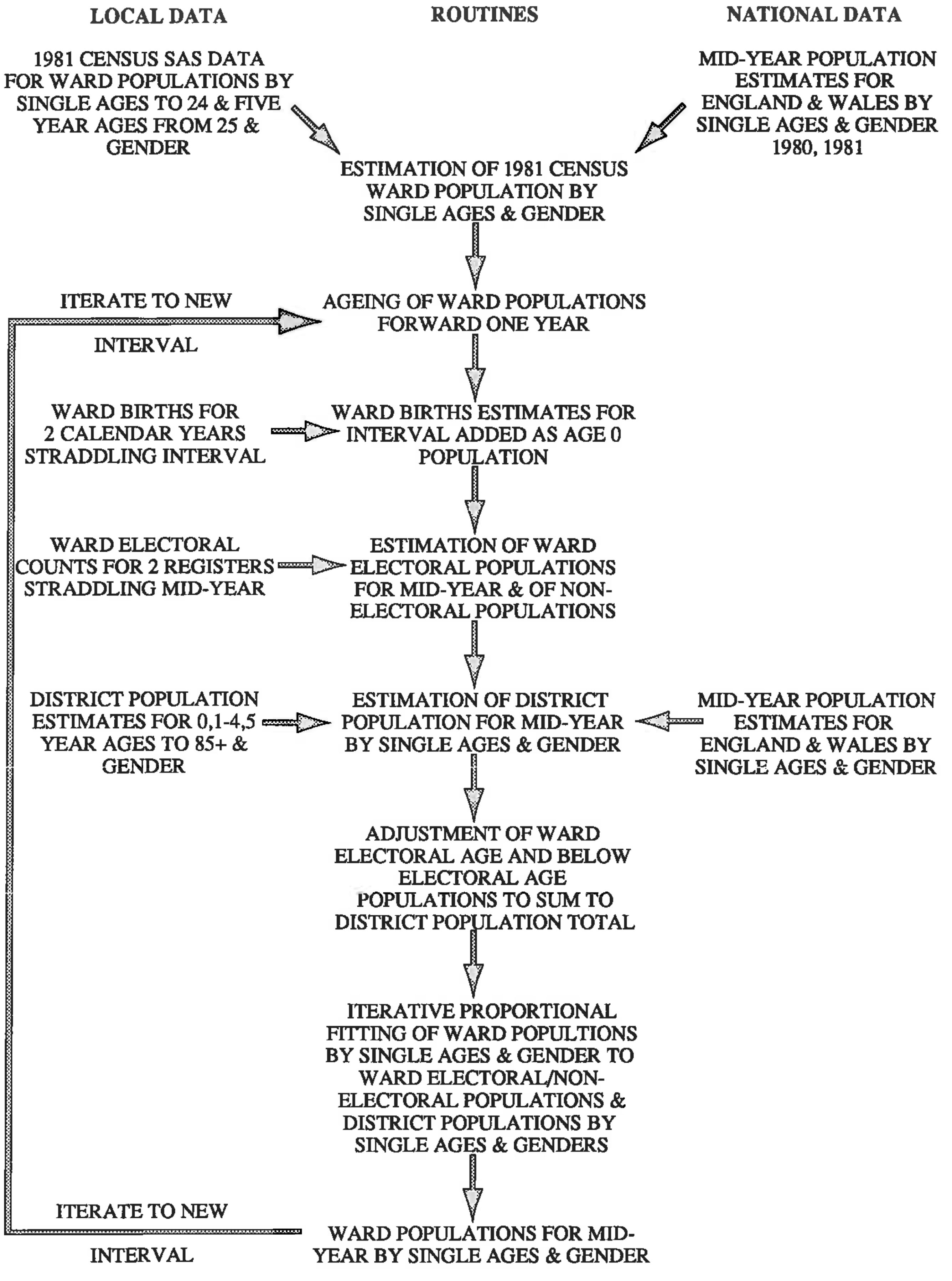

FIGURE 2. **THE STRUCTURE OF THE WARD POPULATION ESTIMATION ROUTINE**

TABLE 2. Numbers illustrating iterative proportional fitting applied to ward population estimation

Ages/genders Wards	Ages 18-49 Males	 Females	Ages 50+ Males	 Females	Total electors
North Ward	12	12	12	12	50
South Ward	10	12	8	8	50
District totals	25	25	25	25	100

3.2 The estimation of mortality

The method for estimating ward mortality rates and non-survival probabilities is set out in Figure 3. The estimates proceed year by year from 1981 to 1990 in a loop.

At the start mortality rates for England and Wales by single ages and gender are computed from published data on calendar year deaths and mid-year populations. These rates are then multiplied by the corresponding ward population estimates to produce expectations of ward deaths, which are summed for the broad ages for which statistics are published. The ratios of observed ward deaths to expected ward deaths for broad ages are then used to adjust the England and Wales death rates to produce estimates of ward mortality rates. The method assumes that the age pattern of mortality for England and Wales applies uniformly across the wards.

From the ward mortality rates which refer to period-ages are then computed the non-survival probabilities for period-cohorts and two summary statistics, namely ward life expectancies and standardised mortality ratios (where the level of mortality in England and Wales is set at 100).

Because the annual rates are liable to fluctuate somewhat erratically from year to year at ward scale, five year averages of the rates, probabilities and summary indicators are computed for periods 1981-85 and 1986-90 to serve as more reliable benchmark data for input to the projection model.

3.3 The estimation of fertility

Figure 4 outlines the method for estimating ward fertility rates, which uses a similar method to that for mortality, although far fewer fertility rates are required. The computation proceeds year by year from 1981 to 1990 in a loop.

At the start fertility rates for England and Wales by five year maternal ages (15-19 to 45-49) are computed from published data on calendar year births and mid-year female populations. These rates are then multiplied by the corresponding ward population estimates to produce expectations of ward births, which are summed to yield total ward births for which statistics are published. The ratios of observed ward births to expected ward births are then used to adjust the England and Wales fertility rates to produce estimates of ward fertility rates.

From the ward age-specific fertility rates are then computed total period fertility rates (measured in children per woman) and standardised fertility ratios (where the level of fertility in England and Wales is set at 100).

Because annual rates are liable to fluctuate somewhat erratically from year to year at ward scale, five year averages of the rates and summary indicators are computed for periods 1981-85 and 1986-90 to serve as more reliable benchmark data for input to the projection model.

3.4 The estimation of emigration

Figure 5 displays the structure of the routines used to estimate ward emigration rates. The data available for estimating emigration rates are not very reliable but it is important to represent every component of change in a properly specified population projection model. The computation proceeds year by a year from 1981 to 1990 in a loop.

At the start emigration rates by single years of age and gender are computed for the United Kingdom using data on emigration from the UK by broad age and gender (derived from OPCS's International Passenger Survey), model migration rates for

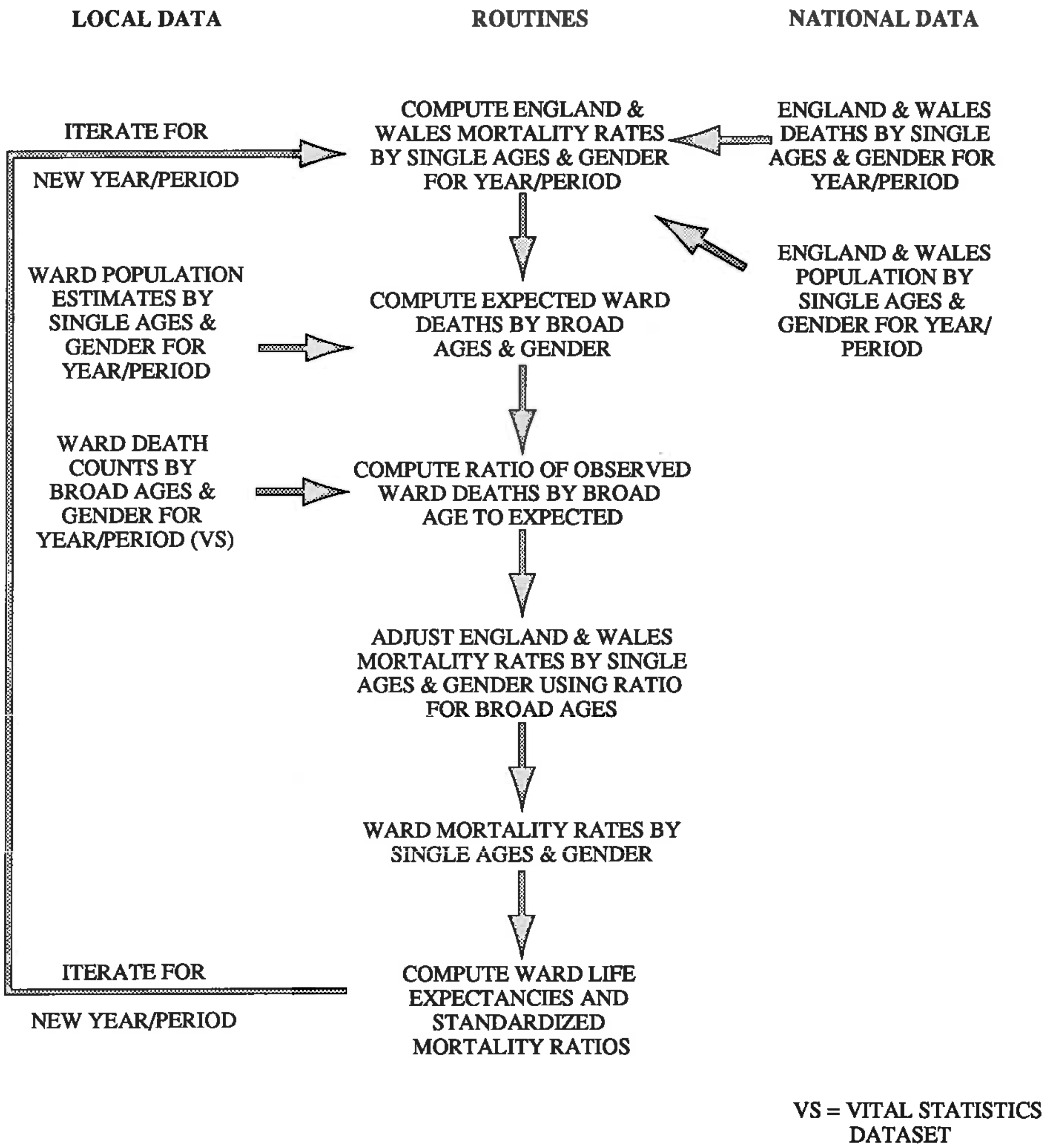

FIGURE 3. THE STRUCTURE OF THE MORTALITY ESTIMATION ROUTINE

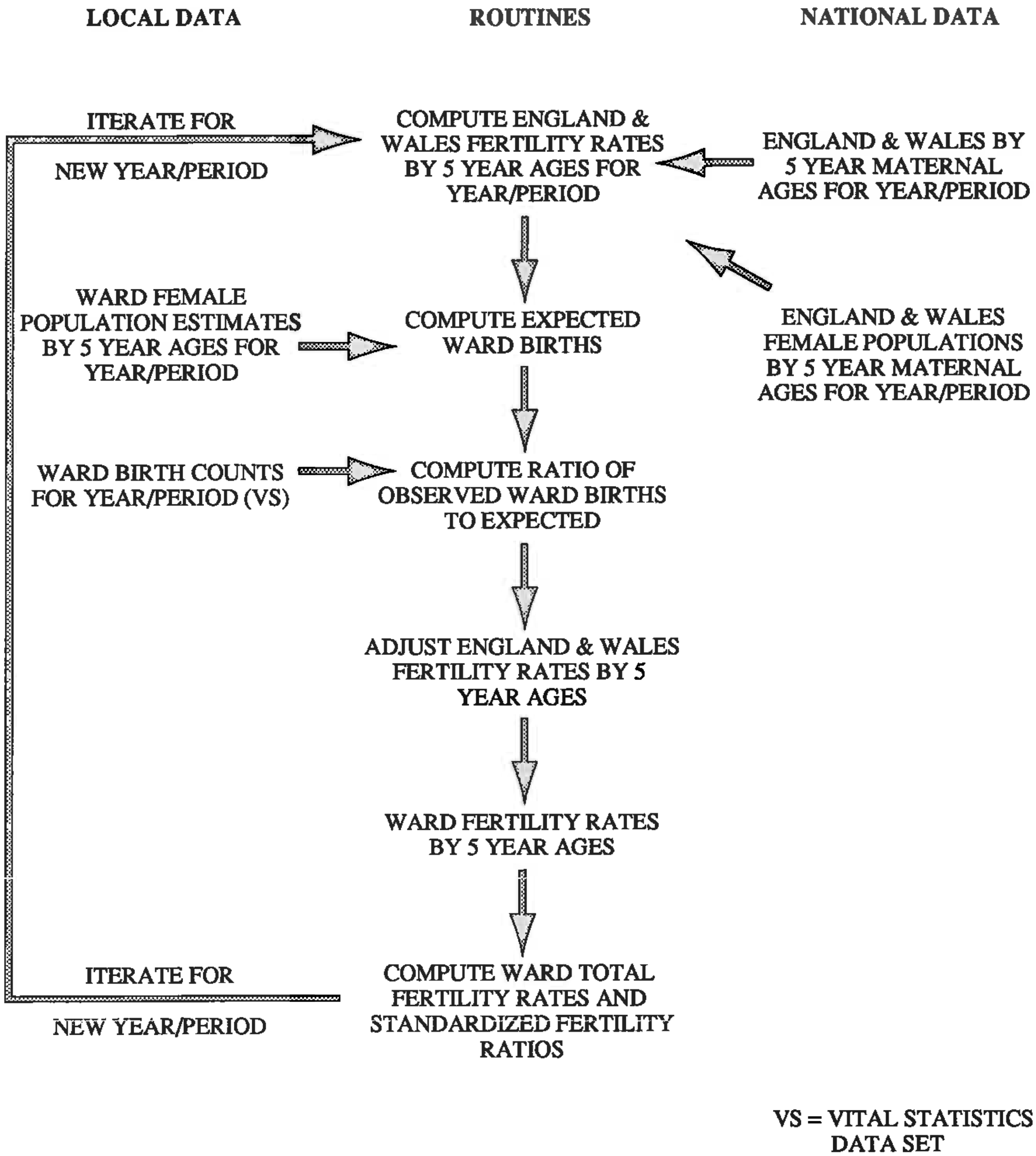

FIGURE 4. **THE STRUCTURE OF THE FERTILITY ESTIMATION ROUTINE**

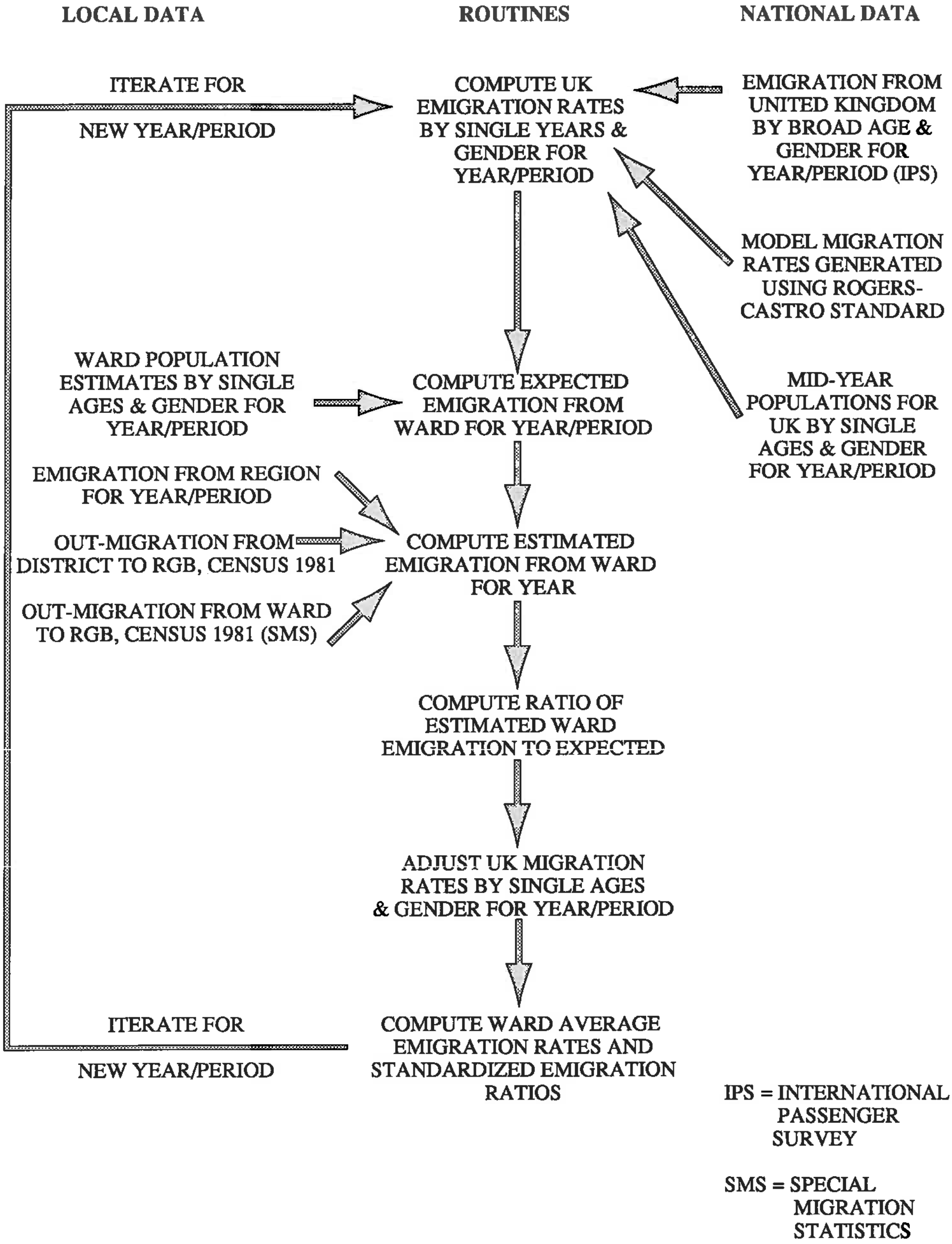

FIGURE 5. THE STUCTURE OF THE EMIGRATION ESTIMATION ROUTINE

single ages generated from the standard defined by Rogers and Castro (1981) and mid-year populations for the UK by single ages and gender.

These rates are multiplied by the ward population estimates by single age and gender and summed over age and gender to produce the expected level of emigration from the ward. This is compared with an estimated level of migration generated using the regional emigration figures for the year factored down employing the ratio of ward out-migration to the rest of GB to district and region sums derived from the 1981 Census migration statistics.

The ratio of estimated emigration to expected emigration is then used to adjust the UK emigration rates to produce ward estimates of emigration rates by single ages and gender. From these emigration rates for period-ages are computed ward emigration and survival probabilities for period-cohorts needed for input to the projection model.

From the ward rates are computed the summary indicators of average emigration rates and standardised emigration ratios (where the level of emigration in the UK is set at 100).

Because annual emigration rates are liable to fluctuate from year to year, five year averages for the rates, probabilities and indicators are computed for periods 1981-85 and 1986-90 to serve as more reliable benchmark data for input to the projection model.

3.5. The estimation of immigration

Figure 6 sets out the method used to estimate immigration rates and immigration and survival probabilities for wards. The method parallels that for emigration in all respects except that information on immigration to the region and to the ward from the 1981 Census is used to factor down annual estimates for the region (Yorkshire and Humberside). Immigration flow estimates are prepared by multiplying the immigration rates by the ward populations.

4 THE ESTIMATION OF MODEL INPUTS (2)

4.1 The framework for internal migration

The framework developed for handling internal migration within the project ion model attempts to marry two principles.

(1) A projection model should capture the influence on population change in an area of events and changes throughout the system.
(2) A projection model should contain a parsimonious representation of the system being modelled, with reliably measured variables.

The first principle suggests that we need to adopt multiregional methods of population projection. The second indicates that a fully explicit multiregional model is not feasible if the system consists of very many areas and a clot of age gender groups. The framework adopted here is therefore a compromise which

(1) uses an aggregate multiregional model to project migration at a broad spatial scale (districts/rest of UK)
(2) uses a set of spatial interaction models to project migration at a fine spatial scale (wards) and
(3) applies methods of age/gender disaggregation only to summations of all internal out- and in-migration.

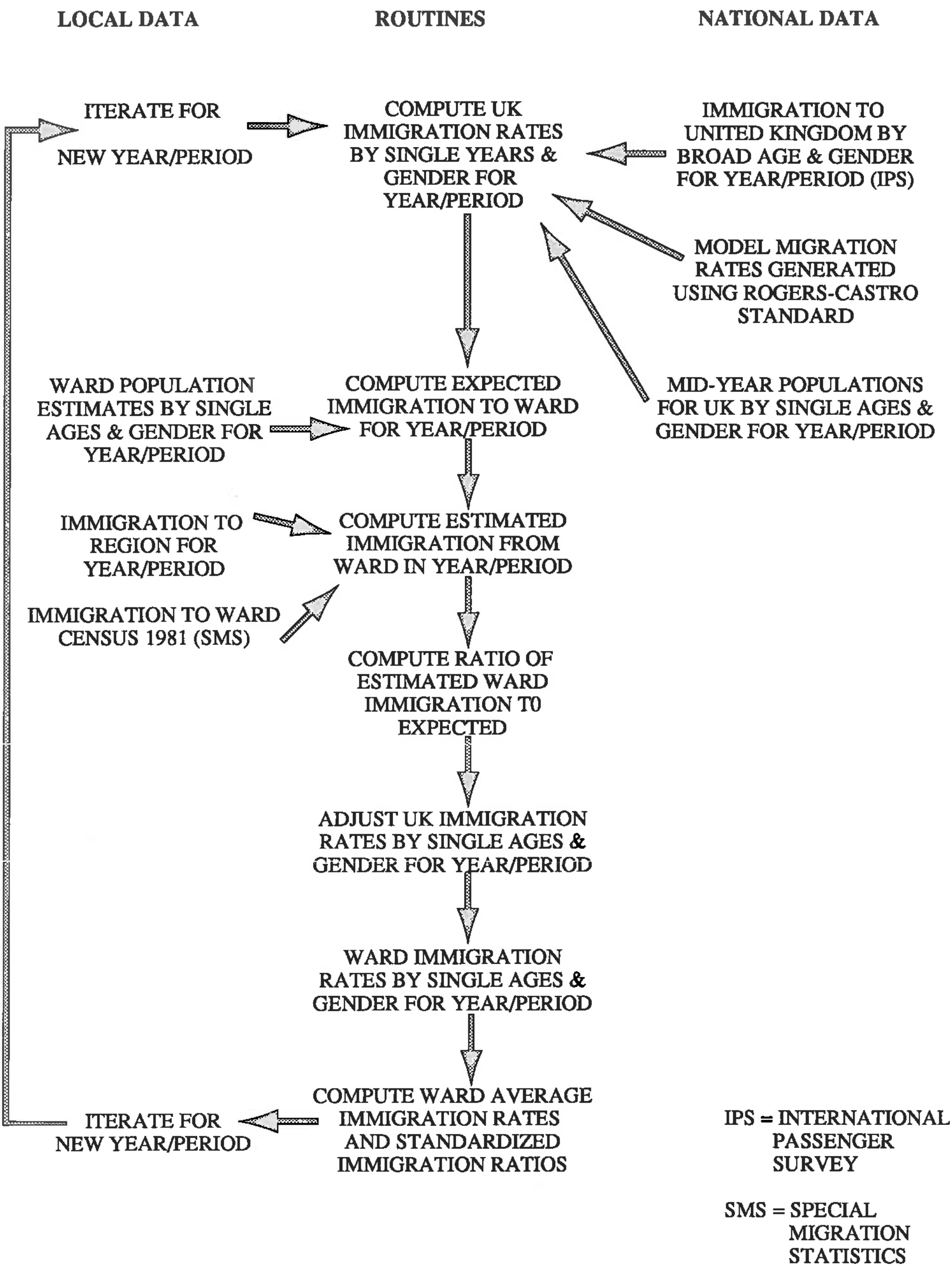

FIGURE 6. THE STRUCTURE OF THE IMMIGRATION ESTIMATION ROUTINE

Figure 7 shows the migration streams (inter-area flows) that are explicitly recognised between the areas for which projections are developed.

At the broad spatial scale the system consists of the district of interest (in this application either Bradford or Calderdale or Kirklees) the districts which interact intensively with that of interest (in this application, the other districts of the former West Yorkshire Metropolitan County) together with a rest-of-the-country zone (which in this application is the United Kingdom less the five districts of Bradford, Calderdale, Kirklees, Leeds and Wakefield).

A simple multiregional model is used at this scale:

migration flow between district i and district j	=	probability of migration from district i to district j	x	start population of district i	(4)

and for the district of interest a ward share is applied to estimate migration into the ward from other districts and out of the ward to other districts.

At the fine scale migration flows within the district of interest are projected using a set of spatial interaction models (detailed later in section 4.3) of the general form:

migration flow from ward i to ward j	=	(balancing factor)	x	(origin term for ward i)	x	(destination term for ward j)	(5)
			x	some function of the cost of interaction between wards i and j (the impedance term)			

The balancing factor is a variable computed internally to the model that ensures consistency. The origin term for ward i is usually the projection of the total number of intra-district migrants leaving ward i. The destination term for ward j is some measure of the attractiveness of ward j to migrants.

The results of the fine scale and broad scale models are summed to give in- and out-migration totals for wards. These are then disaggregated by age and gender using techniques very similar to those employed for mortality, fertility, emigration and immigration (detailed in section 4.4).

4.2 The estimation of inter-district migration

Two sets of data are available for estimating migration flows between districts. The first is the decimal census which provides information on migration flows down to ward scale but for only one year in ten; the second is the National Health Service Central Register (NHSCR) which provides information on migration flows at the scale of Shire country, former metropolitan district and grouping of London boroughs (the former Family Practitioner Committee (FPC)/current Family Health Service Authority (FHSA) areas). Use of the two datasets in combination enables us to estimate migration at the inter-district scale.

The method is outlined in Figure 8. Some initial inputs and computations are made and then the computation proceeds year by year in a loop from 1980-81 to 1989-90. At the start migration flows from the Census and the NHSCR datasets are input for 1980-81 and ward shares of out- and in-migration for the district of interest are computed. Then for each subsequent year migration flows are input from the NHSCR dataset and the ratios to the 1980/81 NHSCR flows are computed. These provide a

DESTINATIONS

Origins		District of interest Wards	WEST YORKSHIRE DISTRICTS (1)	(2)	(3)	(4)	(RUK)
District of Interest							
West Yorkshire Districts	(1)						
	(2)						
	(3)						
	(4)						
Rest of UK							

Zone lists for participating districts:

<u>Bradford</u>		<u>Calderdale</u>		<u>Kirklees</u>	
1	Baildon	1	Brighouse	1	Thornhill
⋮	⋮	⋮	⋮	⋮	⋮
30	Wyke	18	Warley	24	Mirfield
31	Calderdale	19	Bradford	25	Bradford
32	Kirklees	20	Kirklees	26	Calderdale
33	Leeds	21	Leeds	27	Leeds
34	Wakefield	22	Wakefield	28	Wakefield
35	Rest of UK	23	Rest of UK	29	Rest of UK

FIGURE 7. <u>The internal migration streams recognised in the projection model</u>

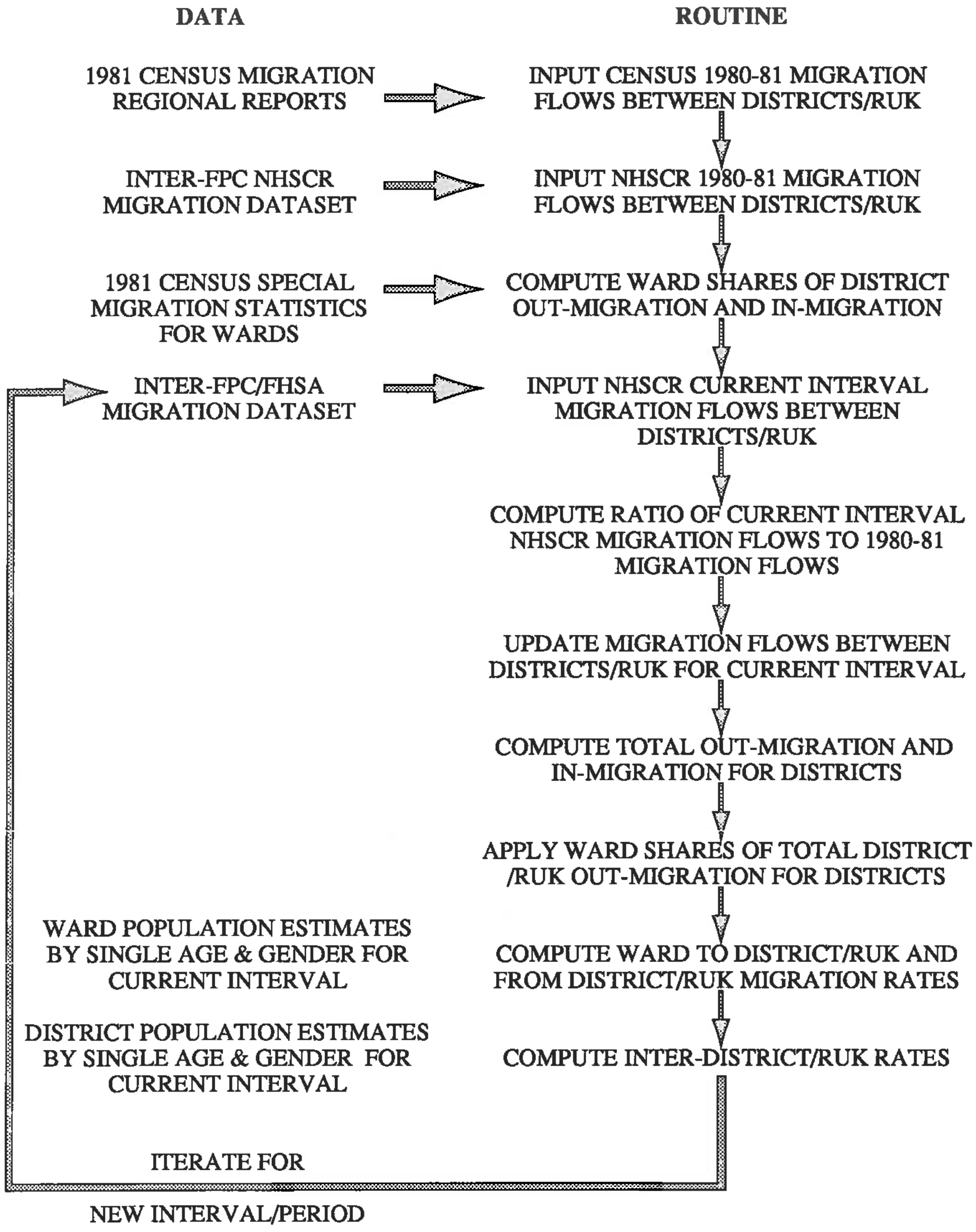

FIGURE 8. THE ROUTINES FOR ESTIMATING INTER-DISTRICT MIGRATION

set of time series indicators for updating the Census migration flows. Note that we use the Census migration data to establish initial levels (and not the NHSCR) because the projection model uses the "transition" concept rather than the "movement" (see section 1.4). The time series ratios are applied to census based migration flows to give an estimate for the current interval (mid-year to mid-year) of inter-district flows. Ward shares are then applied to district total out- and in-migration. The procedures are completed by computing both ward probabilities of inter-district migration for information and district probabilities of inter-district migration for use as benchmark data in the projection.

4.3 The estimation of intra-district migration

Figure 9 sets out the three spatial interaction models (SIMs) used to project intra-district migration between wards. Inter-ward migration can be decomposed into 3 parts which link to housing developments:

inter-ward migration	=	turnover migration (TM) +	new housing migration (HM)	+	migration due to demolition (DM) (6)

Turnover migration and demolition migration are modelled using origin (or production) constrained SIMs. The origin constraints in the TM model are the projected numbers of out-migrants from a ward who go to a ward within the district (including the origin ward). The origin term (not a constraint) in the HM model is the same variable, whereas in the DM variables out-migrants are a function of the numbers of households forced to move.

The projected number of out-migrants (intra-district) is a product of a migration probability multiplied by the origin ward population. Details of the estimation procedures used to update these probabilities are shown in Figure 10 and are similar to those used to produce inter-district migration estimates, except that there is the additional assumption that the inter-district time series is appropriate for updating intra-district migration (not currently captured by the NHSCR).

The destination terms in the TM and DM models are ward attractiveness measures while in the HM model they are the total number of persons moving into new housing in the ward. Initially, the total in-migration flow (intra-district) in 1980-81 can be used to measure revealed attractiveness. However, ward attractiveness may change over time as a result of new housing developments, environmental improvements or changing preferences. The projection model itself can be used to produce new attractiveness factors. Projections are carried out for a single mid-year interval with the best estimate of component inputs for that interval, and the resulting projected populations compared with the independently established population estimates. Attractiveness factors (scaled to unity) are then adjusted in a direction to reduce the discrepancy between projected and estimated populations. The process continues until projected ward populations are reasonably close to estimated.

The final terms in the SIMs are the inter-ward distances (computed from ward centroids) and distance frictions. These are found using the 1980/81 census migration flows and a package for fitting SIMs developed by Stillwell (1984) as in Figure 11. Separate parameters will be determined for the origin constrained models (TM and DM) and the destination constrained model (HM).

4.4 Age and gender disaggregation of internal migration

The final step needed to project internal migration is to disaggregate the flows by age and gender. The method is outlined in Figure 12. It relies on applying model migration rates to projected ward populations and adjusting those rates to yield the out-

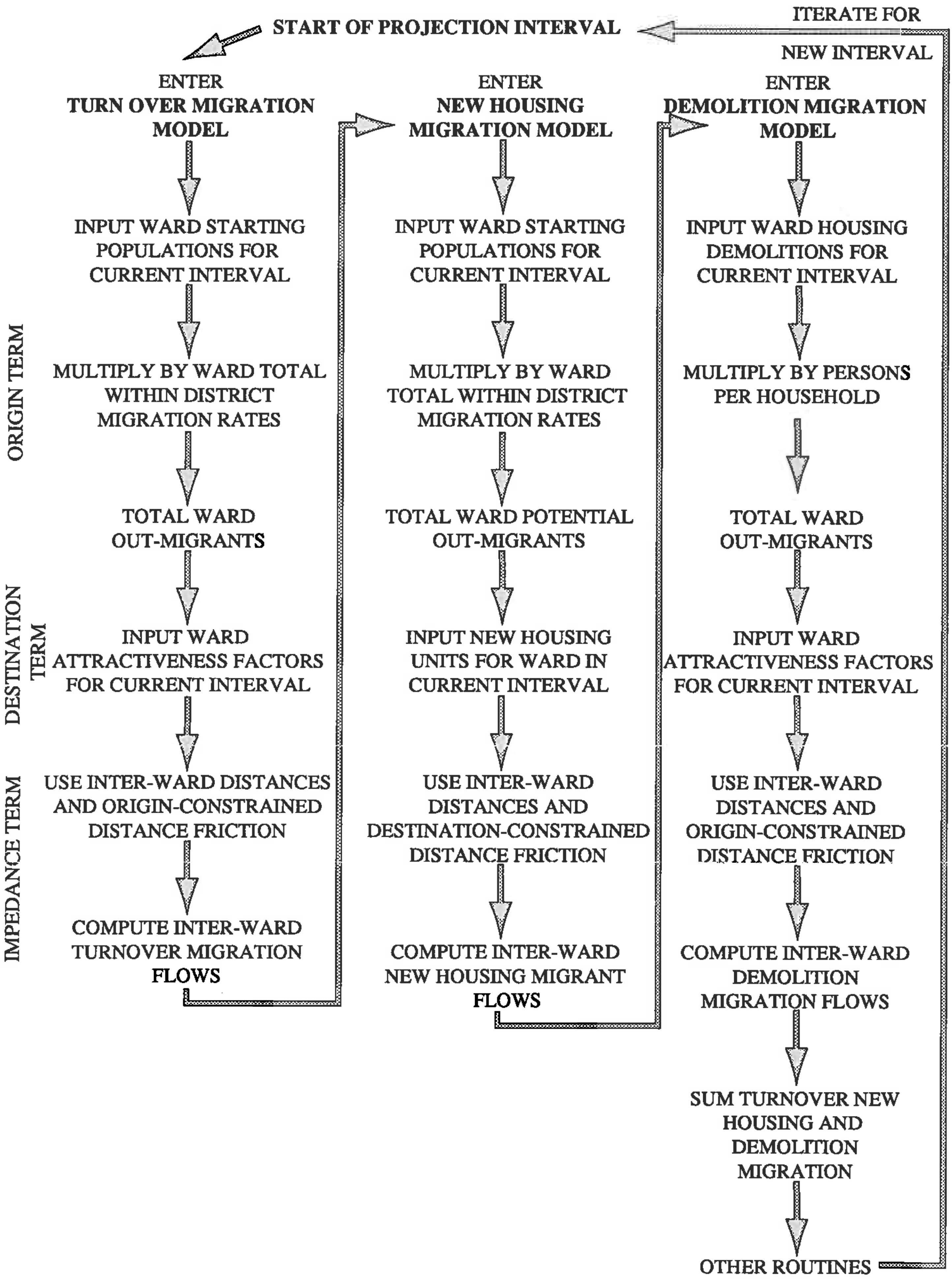

FIGURE 9. THE ROUTINES FOR PROJECTING INTRA-DISTRICT MIGRATION

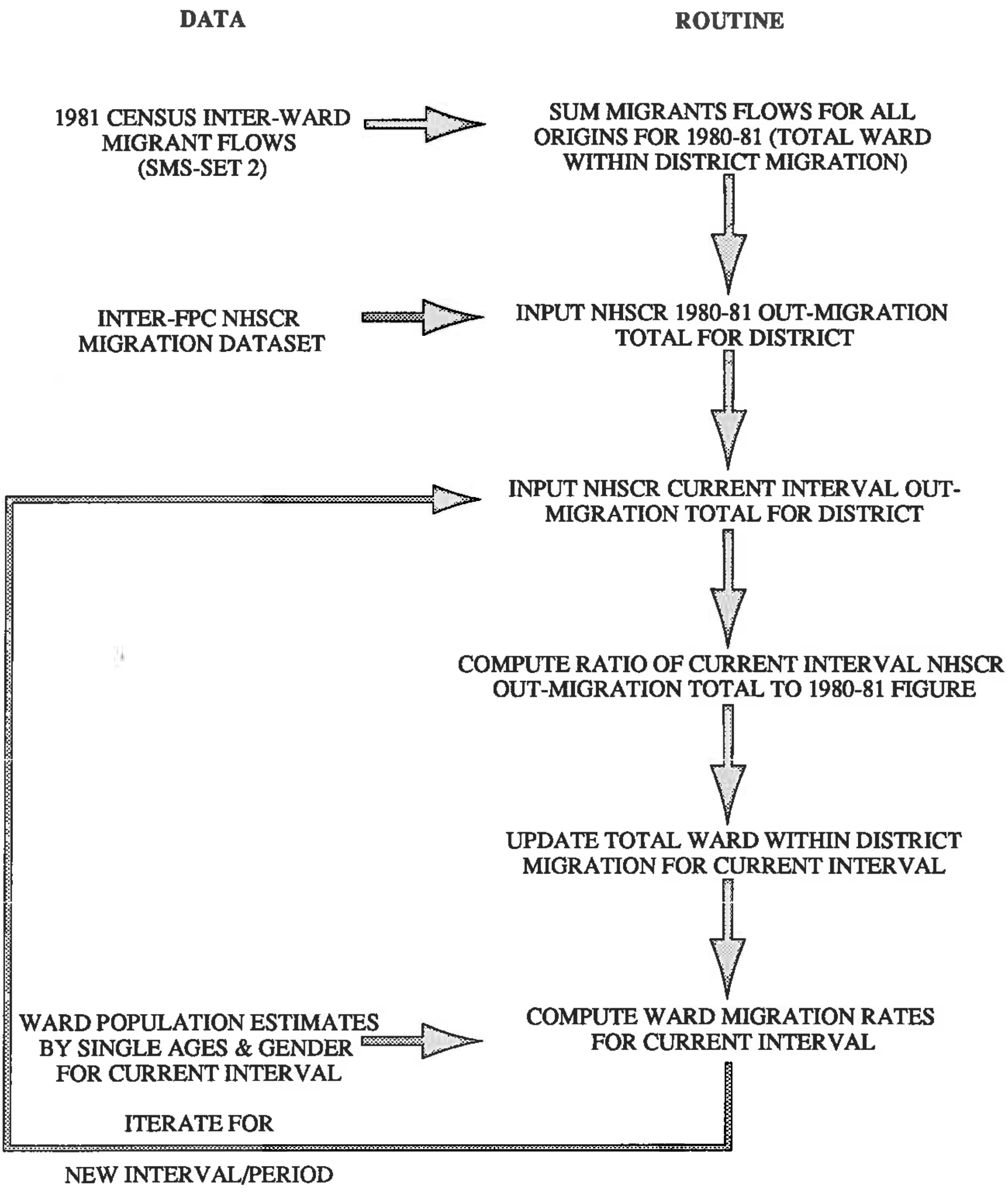

FIGURE 10. **THE ROUTINES FOR ESTIMATING TOTAL WITHIN DISTRICT MIGRATION FOR WARDS**

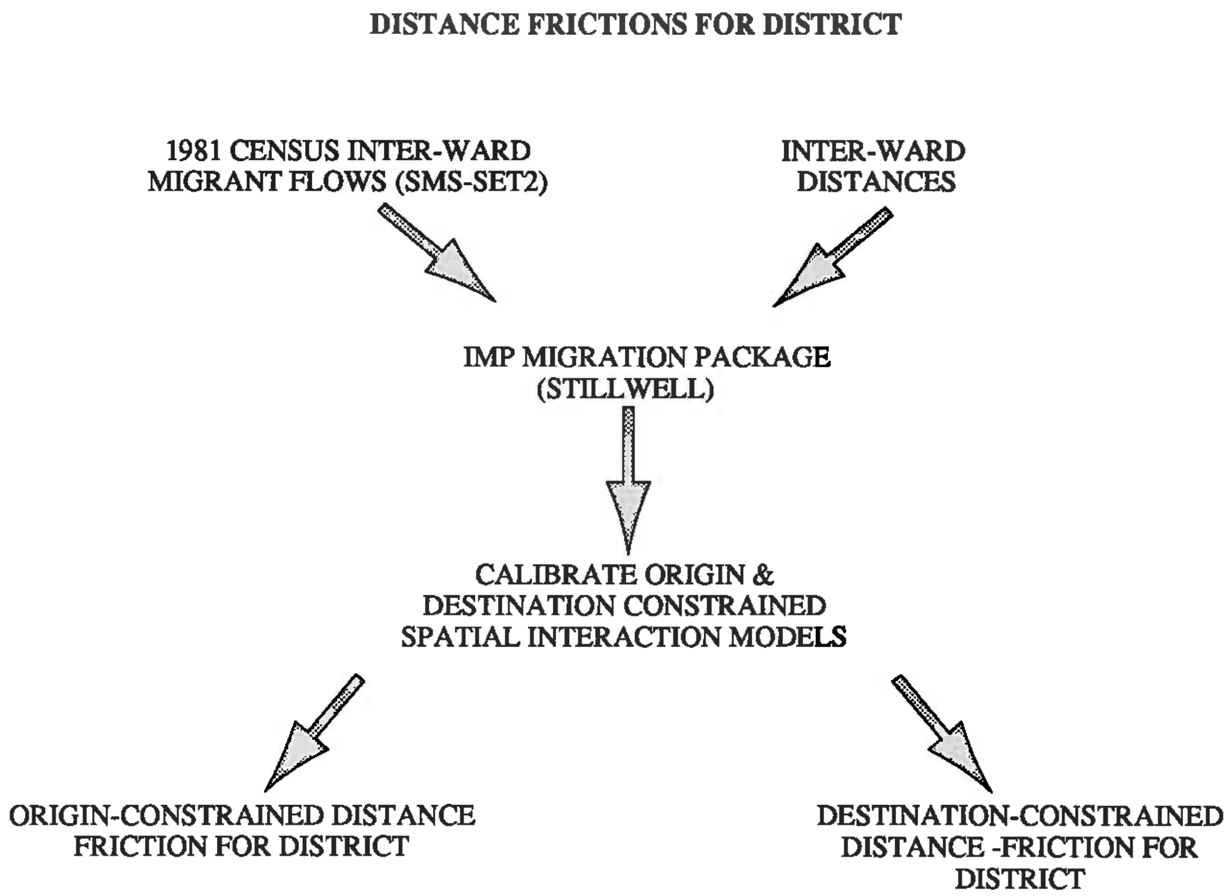

FIGURE 11. THE ROUTINES FOR ESTIMATING DISTANCE FRICTIONS FOR DISTRICTS

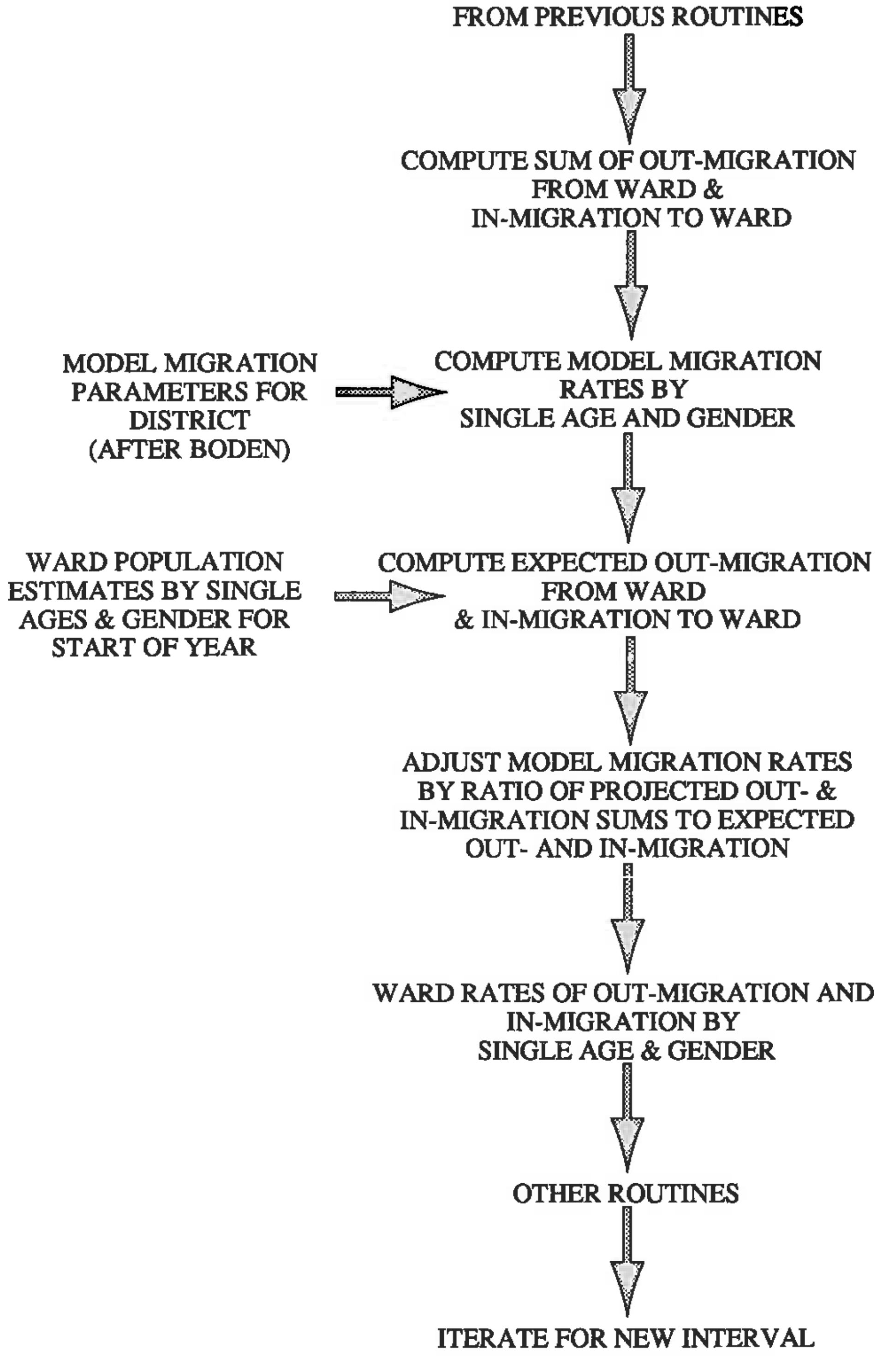

FIGURE 12. THE ROUTINES FOR DISAGGREGATING MIGRATION FLOWS BY AGE AND GENDER

and in-migration totals produced by the aggregate projection of inter-district and intra-district migration. The model migration rates are generated using parameters determined by Boden (1989) for clusters of English and Welsh FPC areas to which the districts of West Yorkshire belong. These parameters generate a characteristic migration curve with an initial peak at young ages, a trough in the mid-teens, a peak in late teens and early twenties and steady decline to later middle and old age, sometimes punctuated by a small rise around retirement (Rogers and Castro 1981).

5. IMPLEMENTATION OF THE MODEL

5.1 Structure of the software

To implement the projection model, a set of programs are being written for implementation on personal computers (IBM PC clones). These programs connect to input data files and generate output files. Programs and data are organised into an information system with the structure shown in Figure 13.

The data files are organised into 4 databases: (1) the raw database containing the input data (mainly as originally published except where changes in ward geography necessitated prior manipulation); (2) the estimate database containing variables, generated by the estimation models, that form the benchmark data for input to the projection model; (3) the scenario database containing a time series of key rates, probabilities, or quantities or key rates of change which are used to modify the benchmark data year by year in the projection; and (4) the projection database containing the principal outputs of the projection model, projected ward populations by single years of age and selected data about the components of change.

Linking data bases are the programs that implement the models described in the body of the paper (the projection model routines and the estimation routines). The scenario design routines are described briefly in the section 5.3.

The implementation software is designed to provide the user with two kinds of functions: carrying out actions to modify the content of the information system and viewing data already in the information system.

The first function provided, shown on the LH side of Figure 13, enables the user to carry out a variety of actions: to add data to the raw database, to update the estimate database using the new data, to design new projection scenarios or to carry out new projections.

5.2 Viewing of model outputs

The second function, shown on the RH side of Figure 13, is the information system which enables the user to access any of the databases and to display their contents in table, map or graph form on the screen and to obtain either paper copies or export file versions of what is shown on the screen.

5.3 Scenario design

Scenario design is the process of determining (guessing?) the trajectories to be followed over the projection horizon by the component inputs. The information used in scenario design must be provided by the user (though the estimate database will be very useful in establishing trends). Full details of the range of methods and therefore choices in scenario design will be provided elsewhere.

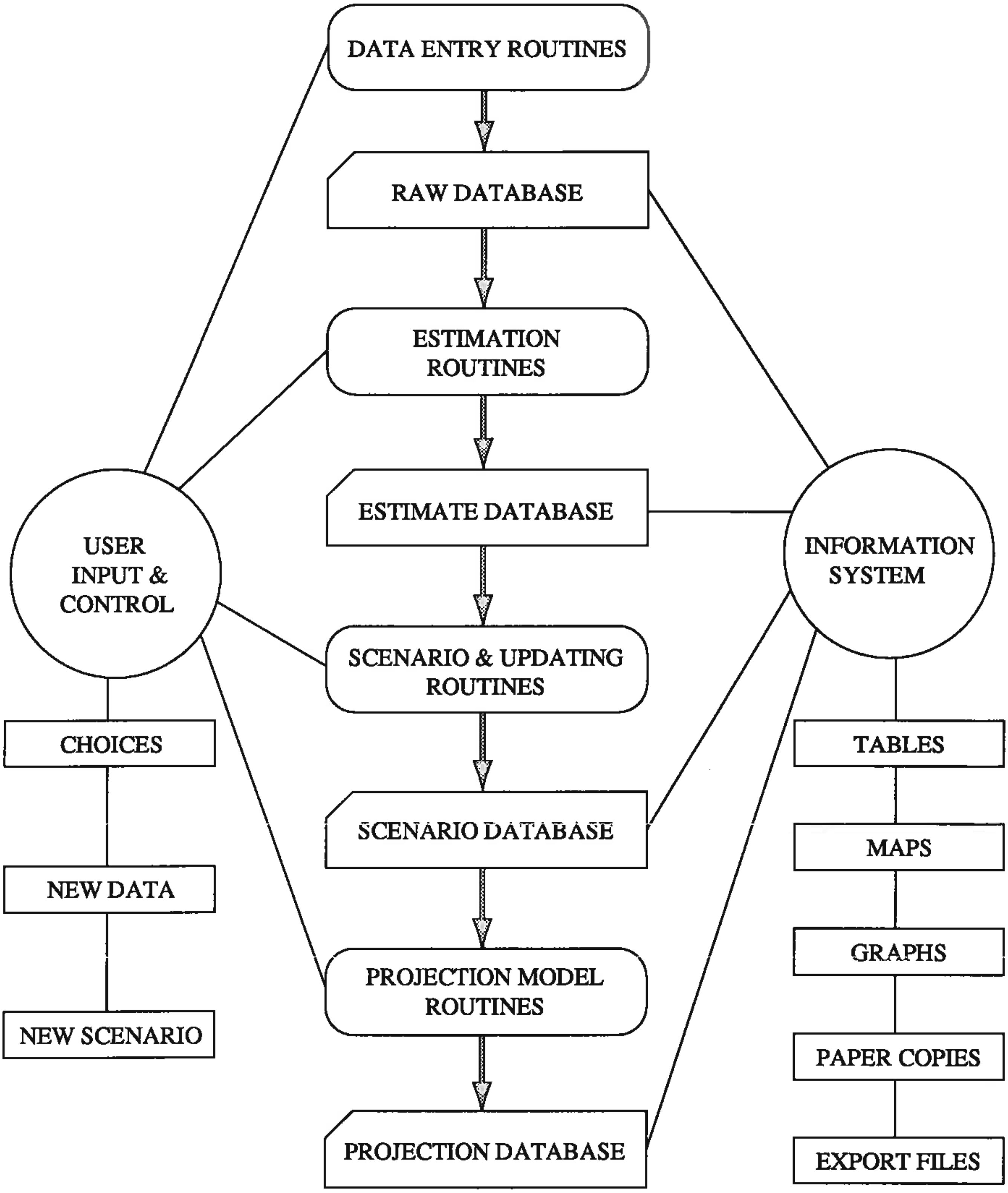

FIGURE 13. THE STRUCTURE OF THE IMPLEMENTATION SOFTWARE

6. FURTHER DEVELOPMENTS

The projection model described in this paper relies, in part, on small area data from the 1981 Census, updated to 1990 using small area, district and national information to produce a variety of benchmark datasets from which projection can begin. The model handles only the demographic characteristics of age and gender, and produces outputs at ward scale only.

Phase (2) of the model will see the development of:

(1) a method for producing flexible geographic output so that the information in the databases can be estimated for any intra-district zones;

(2) a method for converting the projected numbers of individuals into projected numbers of households; and

(3) an adaptation of the projection model for use with different ethnic groups.

REFERENCES

Boden, P., (1989) The analysis of internal migration in the United Kingdom using Census and National Health Service Central Register data. Unpublished PhD thesis, School of Geography, University of Leeds, Leeds

OPCS, (1991) 1989-based population projections for local authority areas in England. Monitor Pp 391/1, Office of Population Censuses and Surveys, London

Rees, P.H., (1989) Old model faces new challenges: a review of the state of the art in multistate population modelling. Working Paper 531, School of Geography, University of Leeds

Rees, P.H., (1990) A manual on subnational population projection. Draft of book in preparation

Rees, P.H., (1991) Method of estimating ward populations for intercensal years. Unpublished paper, School of Geography, University of Leeds

Rees, P.H. and Woods, R.I., (1986) Demographic estimation: problems, methods and examples. In Woods, R.I. and Rees, P.H., (eds.) Population structures and models. Allen and Unwin, London. Pp 301-343

Rogers, A. and Castro, L., (1981) Model migration schedules. Research Report RR-81-30, International Institute for Applied Systems Analysis, Laxemburg, Austria

Stillwell, J., (1984) IMP2: a computer program for inter-area migration analysis and projection - user's manual. Computer Manual 12, School of Geography, University of Leeds

Whitehead, K. and Dugmore, K., (1990) Local and small area projections. In British Society for Population Studies (ed.), Population projections: trends, methods and uses. Occasional Paper 38, OPCS, London. Pp 65-75