COUNTING PEOPLE: PAST, PRESENT AND FUTURE

Philip Rees

WORKING PAPER 93/3

SCHOOL OF GEOGRAPHY • UNIVERSITY OF LEEDS

Inaugural lecture presented at the University of Leeds
on Monday, 15 February 1993 at 5.30 pm
in the Rupert Beckett Lecture Theatre

WORKING PAPER 93/3

COUNTING PEOPLE: PAST, PRESENT AND FUTURE

Philip Rees

School of Geography
University of Leeds
LEEDS
LS2 9JT
United Kingdom

Inaugural lecture presented at the University of Leeds
on Monday, 15 February 1993 at 5.30pm
in the Rupert Beckett Lecture Theatre

CONTENTS

ABSTRACT

The paper examines the procedures used for counting, estimating and forecasting the numbers of people in the past, in the present and in the future respectively. Censuses are the main tools for making comprehensive national population counts at periodic intervals. In recent UK Censuses the set of questions asked has remained fairly stable but the range of outputs being currently produced provides the richest information base yet available on the British population. The ways in which this information base is being delivered to the UK academic community are described in the paper. The paper next considers the methods of keeping the main population counts up to date, which are shown to depend critically on the type of migration information available. Each type is associated with a different species of population account. Population accounts provide consistent benchmark data for population projections. The variety of alternative ways is described in which projection models can be specified and their relative merits are evaluated. The importance of careful development of scenarios of the main components of population change is illustrated for UK mortality and migration experience. The consequences of alternative assumptions for the elderly population into the middle of the 21st Century are explored. The paper concludes with a discussion of the task of demographic estimation which is needed to convert population census and component change data into input variables for population analysis. Why is all this population arithmetic important? In the final analysis, because people count.

ACKNOWLEDGMENTS

I would like to thank all those people with whom I have collaborated over the years, whose names appear in some of the citations at the end of the paper. Within the University these include Alan Wilson, John Stillwell, Martin Clarke, Mark Birkin, Andrew Convey, Marek Kupiszewski, Peter Boden, Chris Duley, Lazarus Zanamwe, Frank Cudjoe, Paul Williamson and Julia Williams. However, only I can take the responsibility for the current piece. The University and the School of Geography have provided the stimulating, resourced environment in which the research described has been possible. My thanks are due to all my colleagues in this respect, but Maureen Rosindale and Alison Manson have been particularly helpful in producing the current paper. Although the paper treats people as statistics, I know that individuals play vital roles of support and encouragement, none more so than my wife Laura.

LIST OF TABLES

LIST OF FIGURES

Count, to number, sum up: to name the numerals up to: to take into account, reckon as significant or to be recognised, to reckon, esteem, consider

(Chambers Twentieth Century Dictionary, Ed.A.Macdonald)

The topic outlined in this paper involves some very simple notions of population arithmetic, the counting of people and of the events and transitions that change those counts. Counting people is important, to use a recent official slogan, because people count. We cannot make judgments about the state of societies or nations without some basic objective information about the numbers of people in various socio-economic states. Because most people spend most of their lives in a sequence of geographically small action spaces (the daily urban system linking residences and workplaces), we need information about numbers of people for small geographical areas. Hence the need for censuses. In the first section of the paper, an overview of the population census is provided and some of its products illustrated.

However, censuses are taken only at quite long time intervals. They are ten years apart in the United Kingdom, for example. Therefore, we need means of updating the information gathered in the census for geographically defined populations. This is where knowledge of the components of population change must be used to build forward from the last census count. This sounds like a very simple task to accomplish but to do it properly it is necessary to use the proper **population accounting** framework, of which there are several, each applicable to particular information gathering situations. This will form the theme of the second part of the paper, along with a couple of useful graphical tools for population analysis.

Anyone who has used published demographic counts will realise that they are always out of date, and that they say nothing as such about the future evolution of the population. To project counts of population forward in time several ingredients are necessary: a model of the way geographical populations are changing, an idea of what are the driving forces behind change and therefore a set of assumptions about the likely evolution of those driving forces. The subject of **population projection** will be addressed in the **third** part of the paper.

Each of the driving forces behind population projection deserves it own special analysis in a geographic context. In the **fourth** part of this lecture I want to look at the way in which two population change components have changed in the recent past in the United Kingdom, those of **mortality** and of **migration**, focusing on the demographic processes altering population counts in areas.

An alternative focus is to examine in detail particular population groups. One which has excited continual interest is that of the **elderly**, which most of us will join and then leave. The **fifth** section of the paper demonstrates how different might be the future from the past, but also how dependent our forecasts of the numbers of elderly are on trends in mortality and thus on aspects of health and social policy currently being worked out.

One crucial assumption is made in much of the discussion in sections one to five of the paper: that there exist adequate, reliable data for the population being studied and for the components of change. What should we do when faced with inadequate information? I argue in section **six** of the paper that if the problem being studied is important, then whatever data are available should be used, along with the knowledge about the nature of populations, to provide the information required for analysis of historical or future change. This theme of **demographic estimation** is developed in the sixth section of the paper and illustrated at trans-national and sub-national spatial scales.

The paper concludes with a final review of why counting people is an essential activity to be continued in a fair, just and sustainable society.

1. POPULATION CENSUSES

1.1 The origins and key attributes of censuses

All primitive civilisations developed the concepts of number and measure as means of organising daily activities, particularly trading transactions. Quite soon the concept of integer numbers was applied to people. The purpose of these early counts of people in Egypt, Babylon, China or Rome was to make lists of persons who could be taxed or who could be inducted into the military. From the point of view of the individual, it was only necessary to know whether his or her name was on the list and with what information attached and to take the necessary action (eg handing over the tax due or going into hiding). But from the point of view of the owner of the list, it was necessary to count up the number of people represented to find out how big was the potential army or associated tax revenue going to be.

Such counting of people came to be part of the activity known as census taking. The definition of a census is well known

"Census The total process of collecting, compiling and publishing data on the demographic, social and economic situation of all persons in a specific territory at a particular time." (Wilson 1985, p26)

Modern censuses, dating from the 18th Century with the first national counts in Iceland (1703), Sweden (1749), Denmark (1769) and the United States (1790), have largely lost their direct link with taxation, although censuses can still be affected by people's perception that names on census lists will subsequently appear on taxation lists. A probable cause of much of the underenumeration reported for the most recent, 1991, Census in the United Kingdom was the erroneous fear by Community Charge avoiders that their names and addresses would be revealed to the relevant authorities. Fortunately, modern censuses have not required the relocation of residents to their birth place which a Roman census at the start of the Christian era required.

Modern Censuses have four essential features (Wilson 1985, p26): (1) individual enumeration, (2) universality within a precisely defined territory, (3) simultaneity and (4) defined periodicity. A return is expected about individuals (either directly or via a household head), about all individuals living in the country at time of the census, for the same date or small range of dates and at regular intervals. Thus the latest UK Census was held simultaneously on the 21 April 1991 in England, Wales, Scotland and Northern Ireland (though not using exactly the same form), some ten years (and 16 days) after the previous Census in 1981. None of these essential features of censuses is fully achieved, even in the best; for example, the Great Britain Census is estimated as having been 2.3% underenumerated in 1981 and 1.9% underenumerated in 1991 (OPCS 1992a).

1.2 The process of census taking

What happens in a Census is that all households and all communal establishments are required by law to fill in the Census questionnaire and to return it to the Census Office (via an enumerator in the case of British Censuses). Figure 1 displays part of the household form (for the first two persons in the household) used in England which asks questions about the sex, date of birth, marital status and relationship to the head of household (that is, the most basic demographic characteristics) from which age-sex-marital tabulations and household classifications can be derived. It should be stressed at this point that no identifiable information on any person or household is released by the Census Offices

1-3 Name, sex and date of birth of people to be included

Important: please read the notes before answering the questions.
In answering the rest of the questions please include:

- every person who spends census night (21-22 April) in this household, **including anyone staying temporarily.**
- any other people who are usually members of the household but on census night are absent on holiday, at school or college, or for any other reason, even if they are being included on another census form elsewhere.
- anyone who arrives here on Monday 22nd April who was in Great Britain on the Sunday and who has not been included as present on another census form.
- any newly born baby born before the 22nd April, even if still in hospital. If not yet given a name, write BABY and the surname.

Write the names in BLOCK CAPITALS starting with the head or a joint head of household.

	Person No. 1	Person No. 2
Name and surname		
Sex	Male ☐ 1 Female ☐ 2	Male ☐ 1 Female ☐ 2
Date of birth	Day / Month / Year	Day / Month / Year

4 Marital status

On the 21st April what is the person's marital status?
If separated but not divorced, please tick 'Married (first marriage)' or 'Re-married' as appropriate.
Please tick one box.

Person No. 1	Person No. 2
Single (never married) ☐ 1 Married (first marriage) ☐ 2 Re-married ☐ 3 Divorced (decree absolute) ☐ 4 Widowed ☐ 5	Single (never married) ☐ 1 Married (first marriage) ☐ 2 Re-married ☐ 3 Divorced (decree absolute) ☐ 4 Widowed ☐ 5

5 Relationship in household

Please tick the box which indicates the relationship of each person to the person in the first column.
A step child or adopted child should be included as the son or daughter of the step or adoptive parent.
Write in relationship of 'Other relative' — for example, father, daughter-in-law, niece, uncle, cousin.

Write in position in household of an 'Unrelated' person for example, boarder, housekeeper, friend, flatmate, foster child.

Person No. 2 — Relationship to Person No.1:
Husband or wife ☐ 1
Living together as a couple ☐ 2
Son or daughter ☐ 3
Other relative ☐ *please specify*
Unrelated ☐ *please specify*

6 Whereabouts on night of 21-22 April 1991

Please tick the appropriate box to indicate where the person was on the night of 21-22 April 1991.

Person No. 1	Person No. 2
At this address, out on night work or travelling to this address ☐ 0 Elsewhere in England, Scotland or Wales ☐ 1 Outside Great Britain ☐ 2	At this address, out on night work or travelling to this address ☐ 0 Elsewhere in England, Scotland or Wales ☐ 1 Outside Great Britain ☐ 2

7 Usual address

If the person usually lives here, please tick 'This address'. If not, tick 'Elsewhere' and write in the person's usual address.
For students and children away from home during term time, the home address should be taken as the usual address.
For any person who lives away from home for part of the week, the home address should be taken as the usual address.
Any person who is not a permanent member of the household should be asked what he or she considers to be his or her usual address.

Person No. 1	Person No. 2
This address ☐ 1 Elsewhere ☐ If elsewhere, please write the person's usual address and postcode below in BLOCK CAPITALS Post-code	This address ☐ 1 Elsewhere ☐ If elsewhere, please write the person's usual address and postcode below in BLOCK CAPITALS Post-code

8 Term time address of students and schoolchildren

If not a student or schoolchild, please tick first box.
For a student or schoolchild who lives here during term time, tick 'This address'.
If he or she does not live here during term time, tick 'Elsewhere' and write in the current or most recent term time address.

Person No. 1	Person No. 2
Not a student or schoolchild ☐ This address ☐ 1 Elsewhere ☐ If elsewhere, please write the term time address and postcode below in BLOCK CAPITALS Post-code	Not a student or schoolchild ☐ This address ☐ 1 Elsewhere ☐ If elsewhere, please write the term time address and postcode below in BLOCK CAPITALS Post-code

Source: OPCS (1992), p.86. Crown Copyright

Figure 1: Part of the form H for Private Households, 1991 Census, England

for 100 hundred years. The names and addresses of respondents are not held on most census files - just the postcode and other locational codes are recorded.

The processing of census data is a long and complicated task involving a large input of person months and computing power. Figure 2 has the temerity to attempt to sum up the processing system in one diagram. It should be stressed this is a general conceptual diagram rather than a precise representation of any particular census. The people in a defined territory supply the answers to the census questions and these are entered onto the computer as either a household record or a communal establishment record. The record for a household will contain a small number (about 5 in the 1991 Census) of attributes that only apply to the household as a unit and a larger number (about 19 in the 1991 Census) of characteristics that apply to the individual persons in a household. The establishment form has a similar structure but the number of persons for which individual data are obtained is much larger.

Data entry and **editing** of the household and establishment forms were major tasks stretching over the period from April 1991 to December 1991 for the UK Censuses (Clark 1992). One new feature of the 1991 Census in Great Britain was the **imputation** of the

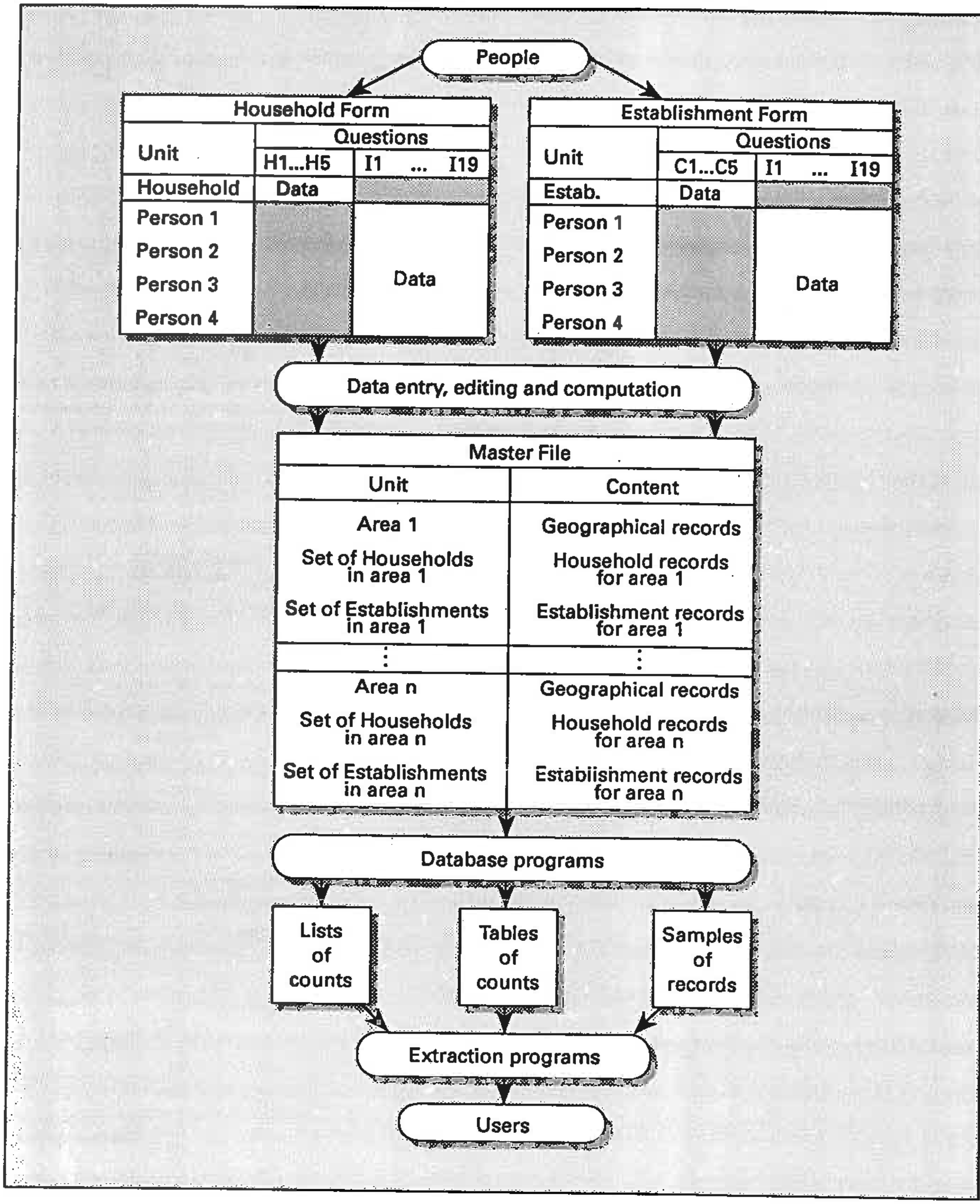

Figure 2: The enumeration and tabulation processes involved in population censuses

characteristics of wholly absent households. These were households that census enumerators, as a result of observation and of talking to neighbours, were certain existed but were away from their usual residence on Census night. They were assigned the attributes of a similar household in the same area in the imputation process. Wholly absent households are an important population group which need to be counted. In 1981 it was estimated that just over one million persons were omitted from the Great Britain Census because they were wholly absent (OPCS 1992a, Annex A, p25). In 1991 more intensive efforts were made to obtain forms from such households on return to their usual residence and 680 thousand persons in enumerated wholly absent households were added along with 869 thousand persons imputed in wholly absent households. These procedures are an example of demographic estimation, a topic which is returned to later in the paper.

The records from the Census will be held in a **Master** file, the general features of which shown in the middle of the diagram. The file is organised by area, with all records for a particular area being listed together, in a hierarchical structure. So there is one set of geographical records for an area, giving all the relevant location codes that can be captured, though not as yet in the UK at a high resolution

(1 metre grid reference). These codes apply to all the households and establishments resident in the area, but do not have to be repeated for each household (that would be wasteful of computer storage). Then there are a set of household records consisting of one set pertaining to the household as a whole, and a second set for each individual in the household. Again the household attributes do not need to be repeated for all individuals. Finally, there are establishment sets of records, each set containing the establishment attributes followed by the individual attributes of establishment residents.

1.3 Census outputs: general nature

This Master file is then read by one of a variety of database programs, which are capable of handling hierarchically organised files, and the Census Offices supply the computer program with instructions for counting up the numbers of individuals or households or establishments that have particular attributes. The output of such programs are either in the form of lists of counts for each area or in the form of tables in which the meaning of the counts is made clear by the "stub" of the tables, that is, the row labels and column headings. Or, the data may be provided as a set of individual or household records. For the first time in UK Census history, some output is being provided from the 1991 Census in anonymous record form (the Samples of Anonymised Records). These records are not copies of parts of the Master Field but carefully screened versions in which classifications have been grouped to preserve confidentiality. This particularly applies to occupational groups of which the census recognises 3,800 categories in the Classification of Occupations and Directory of Occupational Titles (CODOT). These will be grouped so that it will not be possible, for example, to identify unique individuals such as the Bishop resident in the County of Durham, formerly Professor at the University of Leeds.

For the purposes of data transfer (from the Census Office to purchasing user organisations), the output data are normally transferred in list form together with information about the exact definition of each of the counts in the list. End users of Census data, however, will normally be interested in readable tables of population statistics. These can be produced using database or statistical or special purpose packages with tabulation facilities. An example of such a tabulation package is SASPAC91 (the Small Area Statistics Package for the 1991 Census) which enables users to extract tables of counts for geographical areas using SASPAC (1992). This package uses the list form of data prepared by the Census

Offices containing the population counts and supplies a separate file containing the information on row and column labels that converts the "naked" data into readable information.

An example (taken from a Geography class exercise) of a file of instructions that a user might employ with SASPAC91 is as follows

```
* CENSUS1 CMD FILE PREPARED BY G.E.OGRAPHER
INPUT SYSTEM FILE NAME = WYORKWLH
INPUT FRAMEWORK FILE NAME = TLBS132
INCLUDE 08DAFJ
PRINT TABLE L51
OUTPUT PRINT FILE NAME = CENSUS1
END
FINISH
```

Note that two input files need to be named: the first contains the count data (WYORKWLH = West Yorkshire, Ward level, Local Base statistics, one Hundred percent datafile) and the second with the table stubs (TLBS 132 - table stub information for the Local Base Statistics suitable for printing on 132 character wide printers). Once users have obtained the counts or manipulated counts via an extraction program, they will usually then carry out their own further analysis using other software (statistical packages, spreadsheets, mapping programs, or their own programs).

1.4 The outputs of the 1991 Censuses in the United Kingdom

While Figure 2 shows in general the ways in which population count data are output from Censuses, it is useful to examine in detail the variety of different sets of statistics that are currently being produced from the 1991 UK Censuses. Here a very brief overview is provided: full details are available in **The User's Guide** to the 1991 Census (Dale and Marsh 1993).

Table 1 lists the main outputs from the 1991 Census in its rows and gives dates for first availability. These dates should be read as a guide only because the timetable of the Census Offices is subject to revision in the event of processing delays. The columns of the table refer to the three component parts of the UK which have their own Census Office. The Office of Population Censuses and Surveys (OPCS), responsible for England and Wales, collaborates closely with the General Register Office Scotland (GRO(S)) to produce many publications and datasets for Great Britain, while the Census Office Northern Ireland operates under different legislation and produces a different range of products, some of which parallel those in Great Britain.

The table is divided into two panels according to the medium in which the data are delivered. The items in

Table 1 The principal outputs from the 1991 Censuses of Population for England and Wales, Scotland and Northern Ireland

Outputs	Months of availability		
	England & Wales	Scotland	Northern Ireland
PRINTED DOCUMENTS, AVAILABLE IN LIBRARIES			
National Summary Reports, Part 1	(GB) 03/93		11/92
National Summary Reports, Part 2	(GB) 08/93		n/a
National Summary Monitors	(GB) 12/92		n/a
Topic Statistics Reports	(mostly GB) 01/93-04/93		01/93-12/93
County Reports, Part 1	07/92-02/93		n/a
Scottish Region Reports, Part 1		07/92-02/93	
County Reports, Part 2	01/93-06/93		n/a
Scottish Region Reports, Part 2		01/93-06/93	
County Monitors	07/92-10/92		n/a
Scottish Region Monitors		07/92-10/92	
MACHINE READABLE STATISTICS			
County Monitors	12/92		n/a
Scottish Region Monitors		12/92	
Local Base Statistics & Small Area Statistics, 100%	07/92-01/93	09/92-01/93	12/93
Local Base Statistics & Small Area Statistics, 10%	01/93-04/93	01/93-04/93	12/93
ED to PC Directory	07/92-12/92	n/a	12/93
Special Migration Statistics	(GB) 07/93		n/a
Special Workplace Statistics	(GB) 11/93		
Samples of Anonymised Records (2%I,1%H)	(GB) 05/93		08/93
Longitudinal Study 1991 Census Update	12/93	n/a	n/a

Source: OPCS (1992), p.20 and HMSO (1992)

Notes:
1. n/a = not applicable
2. The 1993 dates are planned dates

the top panel consist of traditional printed volumes published by Her Majesty's Stationary Office (HMSO), the official government publisher (except for the Monitors which were published by OPCS and GRO(S)). The area reports for counties (England and Wales) and for regions in Scotland are being published first, as each area's Master file is processed; national summary reports and topic volumes for whole national units can only appear when all areas have been processed.

In the bottom panel of Table 1 are listed the datasets which will be made available in machine readable form, though users may request computer-printed output where appropriate. The data are normally written onto half-inch magnetic tape at the Census Offices, dispatched to users who read the data onto computer disk for faster access by extraction software. Selected statistics may become available through third parties or other media (floppy diskettes or compact disks). Full details of all these datasets are provided in the 1991 Census User Guides produced by OPCS and GRO(S), and a summary description is provided in 1991 Census definitions (OPCS 1992b).

1.5 The dissemination of census outputs to the academic community

How are these various 1991 Census outputs being made available to users in academic institutions? Essentially, they are made available free (in cash terms) at the point of use through collective purchase and organisation through the 1991 Census of Population Programme. This Census Programme is financed by the Economic and Social Research Council (ESRC-£1.5 million), the Information Systems Committee of the Universities Funding Council (ISC-£1.5 million) and the Department of Education Northern Ireland (DENI-£0.12 million). ESRC oversees the programme and negotiates contracts with the Census Offices and other data supplying organisations. There are now eight elements to the Census Programme which are laid out in Table 2. Full details of the programme are provided in Rees (1992a). Here some of the key characteristics of the programme are highlighted. The datasets are held and supported centrally by a series of ESRC and ISC funded units (although other institutions can acquire whole national datasets if they wish). Users access the data via the national academic computer network called JANET (Joint Academic NETwork). A set of projects aimed at providing training in the use of Census data is distributed around the country: users can obtain advice from their nearest project. Particular methodological problems are being addressed by a series of development projects at a variety of universities. A seminar programme provides a forum for reporting on the results of this work and in

Table 2 **Arrangements for the maintenance and dissemination of Census data to the UK academic community under the ESRC/UFC-ISC 1991 Census of Population Programme**

	The Components	Location	Access/Content
(i)	the Census datasets	MCC Other Registered Sites	Registered Users Via JANET
(ii)	the Locational reference datasets	MCC Edinburgh Data Library Other Registered Sites	Registered Users Via JANET
(iii)	the Units	CDU, Manchester CMU, Manchester DA, Essex LSSP, City	Via JANET
(iv)	the Training Programme	City/Manchester	User's Guide to the 1991 Census
		Southampton	Tutorial & User Guides
		Leicester/MRRL	Sample Datasets
		Leicester/MRRL	Trainer's Resource Base
		Cardiff	Training Workshops
		Manchester/City	Training Programme
(v)	the Development Programme	Southampton	Spatial Models of Census Data
		Southampton	Aggregation Issues
		Warwick	Occupational Definition
		Manchester	American Census Lessons
		Leeds	Migration Analysis, Population Synthesis
		Leeds	Aggregation Issues
(vi)	the Seminar Programme	Meetings	Census Analysis Group
(vii)	the Research Programme	For 1994-95	To be determined
(viii)	Coordinator	Leeds	To any census user in the academic community

Source: Rees (1992a)

Notes:
MCC = Manchester Computing Centre
CDU = Census Dissemination Unit, MCC
CMU = Census Microdata Unit, University of Manchester
DA = ESRC Data Archive, Essex
LSSP = Longitudinal Study Support Programme, City University
JANET= Joint Academic Network
MRRL = Midlands Regional Research Laboratory

providing a collaborative network. One of the tasks taken on by Census Analysis Group members at Leeds and Newcastle was the creation of a machine readable version of the County Monitor/Scottish Region Monitor data, which provided researchers with a simple national dataset in December 1992.

The Census Programme plays the very important role of negotiating conditions under which the data may be used. All users must agree to and sign a set of conditions before access is granted. Academic users can only use the data for research or teaching purposes. Any commercial use must be negotiated with the supplying Census Office or other agency. The conditions of use attached to the use of the Samples of Anonymised Records are particularly strict; users must agree not to try to match records from the Samples with known individuals or households and not to claim to have successfully done so. A breach by a user of these conditions may result in the withdrawal of access to the Samples of Anonymised Records (SARs) by all users at the same institution.

The SARs are a new dataset which will be of considerable utility to the broad spectrum of social scientists. Each user will have the freedom to tabulate the individual and household records in any way appropriate, rather than being constrained by the prior decisions about tabulation reached by the Census Offices. The SAR for individuals (2% sample) will be available for local government districts with a resident population of 120,000 or more. Some 278 areas have been defined in Great Britain as amalgamations of the 459 districts, while about 10 will be defined for Northern Ireland. This will make the SARs a geographical by rich data resource.

1.5 Some results from the 1991 Census: limiting long-term illness

It is important not to conclude this discussion of counting people through the Census without a brief peek at some early results from the 1991 Census. One of the new questions asked in the 1991 Census was

> 12 long-term illness
> Does the person have any long-term illness, health problem or handicap which limits his/her daily activities or the work he/she can do?
>
> Include problems which are due to old age.
>
> (OPCS, 1992b, p88)

This variable is intended to aid Local Authorities in planning provision of home based services in 1993 under the "Care in the Community" initiative in which overall responsibility for the chronic ill or disabled is transferred to Local Authority Departments of Social Services.

The percentage of the population reporting limiting long-term illness is reported in Table E of the Census Monitors for counties in England and Wales and for regions in Scotland, and for districts in both.

Figure 3 plots these percentages for counties (England and Wales) and Scottish regions. The map pattern resembles that for mortality studies by Rees (1989a) for the 1980-86 period and for earlier periods by Howe (1963). Both the morbidity map and the mortality map show simultaneously (1) a gradient from the South and East (lower reported illness/lower mortality) towards the North and West (higher reported illness/higher mortality), (2) more reported illness and higher mortality in denser urban counties than in the less densely populated areas, and (3) particularly high levels of limiting long-term illness and of mortality from lung conditions (pneumoconiosis, silicosis, emphysema, bronchitis, lung cancer) in coalfield counties (Durham, South Yorkshire, Gwent, Mid Glamorgan and West Glamorgan).

However, the relationship between limiting long-term illness and mortality is far from perfect. Figure 4 graphs the census variable against the standardised mortality ratio (SMR) for all ages over the calendar years 1980 to 1986, computed using a UK benchmark (Rees 1989). The Island areas of Scotland and Inner and Outer London are grouped, and the Northern Ireland

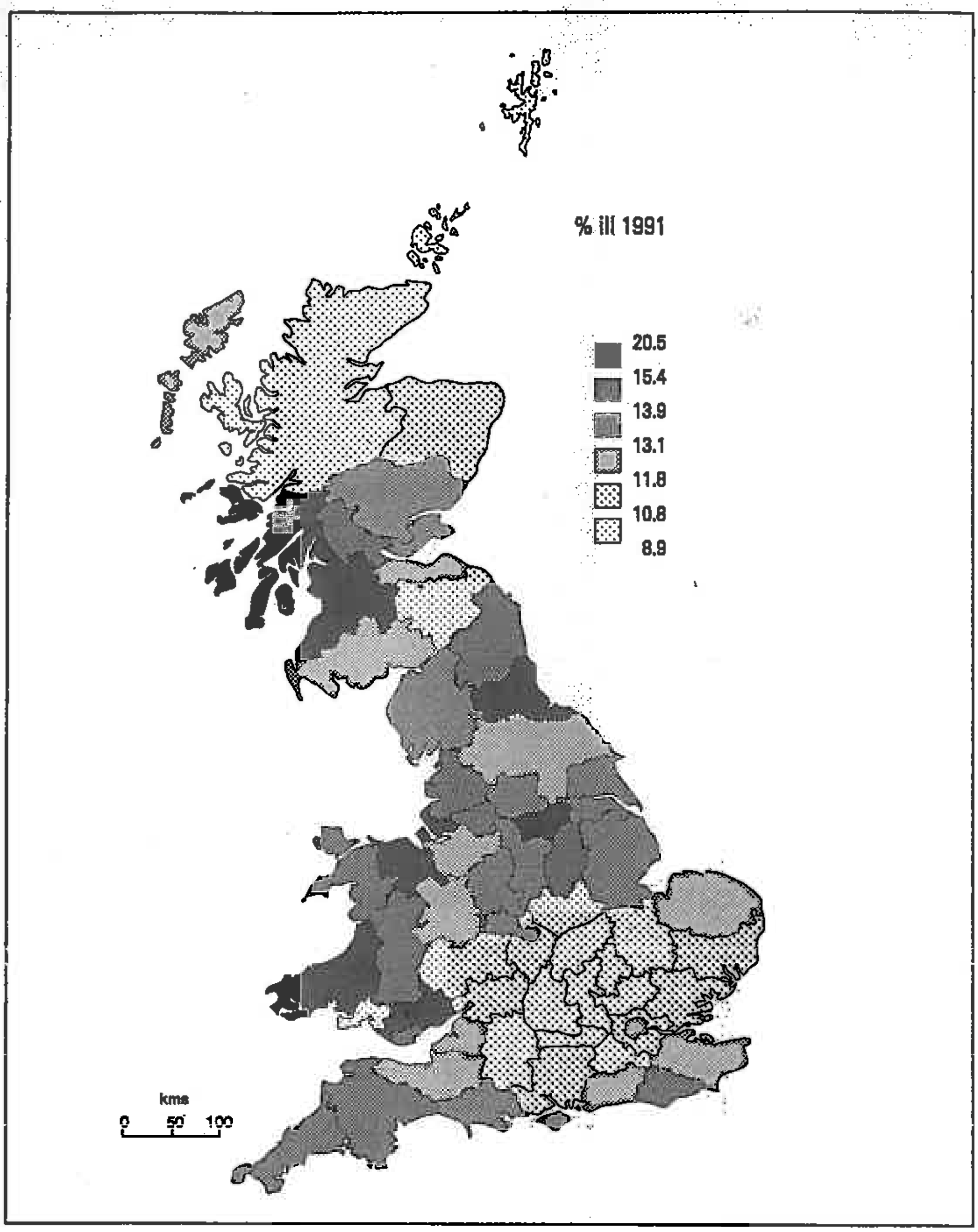

Figure 3: Percentage reporting limiting long-term illness, GB counties/regions, 1991 Census

Source: OPCS County Monitors and GRO(S) Scottish Region Monitors. 1991 Census. Crown Copyright

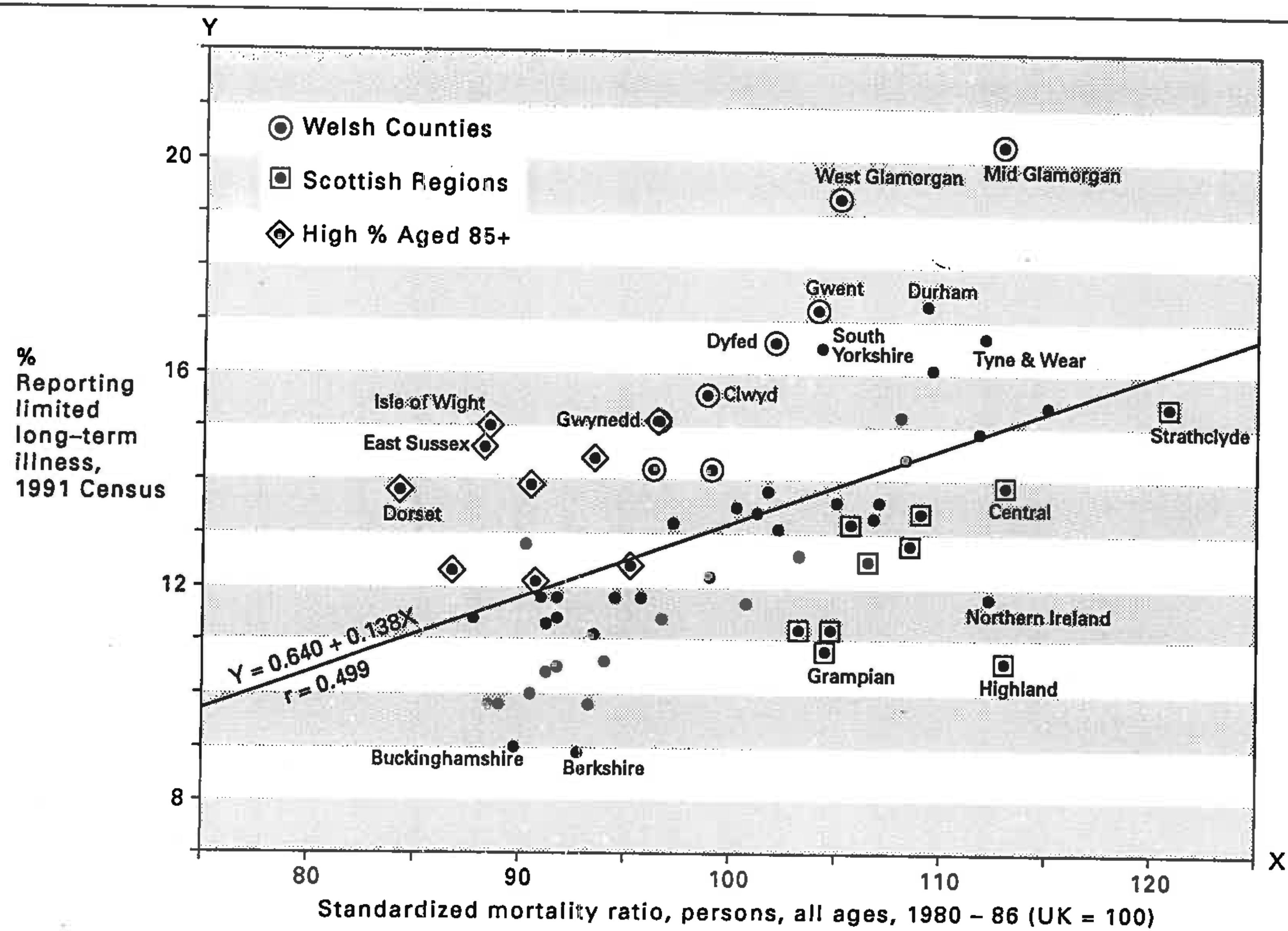

Figure 4: The relationship between limiting long-term illness and mortality

percentage and SMR are included as well. There is a weak positive relationship but only 25% of the variance in limiting long-term illness is associated with the SMR.

Some clues as to other important variables influencing the pattern of limiting long-term illness can be gleaned firstly by looking at the nature of the areas departing strongly from the average relationship and secondly by looking at the illness map at a finer spatial scale (Figure 5). The district map emphasises the importance of coal field locations; a history of employment in coal mining clearly leaves a severe disability legacy. Also picked out are districts around the South coast and East coast of England and on the North coast of Wales which contain high concentrations of the very elderly (aged 85 and over) who are more likely to suffer long-term illness.

There also appears to be a cultural factor at work. All Scottish regions and Northern Ireland, with high mortality experience, lie below the average relationship (regression line) while all Welsh counties are displaced well above the line. The Census variable is a self reported and therefore subjective measure and it looks likely that Welsh and Scott react somewhat differently to their health status.

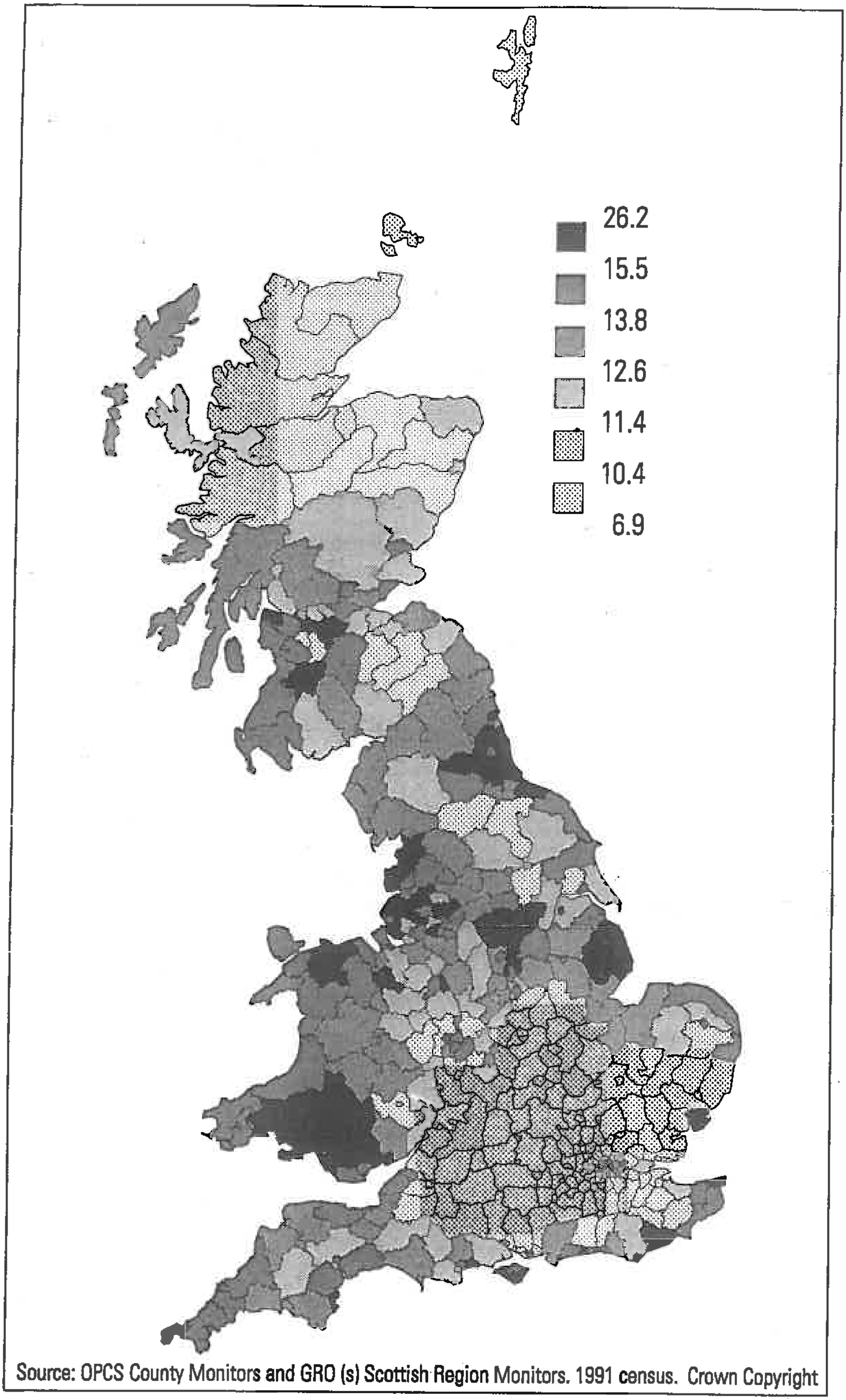

Figure 5: Percentage of the population reporting limiting long–term illness, districts in Great Britain, 1991 Census

The relationship between morbidity and mortality is one not just of academic interest. The level of mortality in a health district is one of the elements that feed into the Resource Allocation formulae for distributing National Health Service funds. But the NHS is not a burial service, it is a service for people in ill-health. However, before we substitute the census variable for the mortality statistic in the funding calculation, we need to understand very clearly what has been measured. Further analysis using the much more detailed statistics available from the Local Base Statistics (LBS) and Small Area Statistics (SAS) is therefore indicated. Table 3 lists the tables that contain information on long-term illness. For example, Tables L12, L13 and L35 (which record the age, sex and marital status of all household residents) enable us to examine for 10,000 spatial units (wards in England and Wales, postcode sectors in Scotland) the relationship between the demographic characteristics of age, sex and marital status and long-term illness. A similar analysis for 145,000 spatial units can be carried out using Tables S12, S13 and S35 (in a more coarse fashion). This is a project that would fit very well into the proposed Census Research Programme.

Table 3 Tables giving information on limiting long-term illness in the 1991 Census LBS and SAS

LBS	SAS	Title
L04		Medical care establishments
L12	S12	Long-term illness in households
L13	S13	Long-term illness in communal establishments
L14	S14	Long-term illness and economic position
L18		Imputed residents
L29	S29	Dependants and long-term illness
L44	S44	Household composition and long-term illness
L47	S47	Households with pensioners: housing
L49		Ethnic group: housing
L68	S68	Floor level of accommodation
L75	S75	Hours worked (10% sample)
L86	S86	SEG of households and families (10% sample)

2. POPULATION ACCOUNTING

Inevitably, Census information becomes dated. Some fourteen months elapsed between Census Date and the release of the first LBS/SAS dataset for the Isle of Wight and some thirty two months will have elapsed before the last output listed in Table 1 will have become available to academic users. In many European counties (Rees and Willekens 1986) population registers are maintained which record the current location and demographic characteristics of each member of the population. However, the majority of countries, including the United Kingdom, do not keep compulsory and comprehensive registers (except in times of national emergency), and must therefore rely on the information generated by partial registers and surveys to make estimates of the population count since the last census. To make such estimates it is necessary to have a structured view about the demographic sources of population change and to organise the components of change within a suitable framework. One such framework is that of population accounting.

2.1 Population accounts defined

Population accounts are tables of population flows that link together population stocks at two points in time. The pioneer in developing demographic accounting was the British economist Richard Stone (eg Stone 1971; UN 1976), who saw them as part of a much wider system of social and demographic monitoring of the activities of national populations.

Drawing on his work in developing national economic accounts, Stone distinguished two kinds of population account: the **open** account which shows how populations over two successive periods are linked, and the **closed** account which reveals the population flows within a period which help the starting population stock to evolve to become the ending population stock. Table 4 gives a simple illustration for the United Kingdom in the style of Stone's open account.

Stone's open account concentrates all transitions (changes in state) in the survivors who move from being the end population of a previous period to being the start population of the next. This is a convenient conceptual position when studying graded educational systems when the shift from one grade to another takes place between the end of one educational year and the start of the next. However, in more general situations such grade transitions can occur at any time during a time interval, and are represented by a closed account. The remainder of the discussion will focus on this sub-species, sequences of which can always be summarised and re-represented in open form.

Table 4 **An open demographic account for the United Kingdom population, 1985-90**

Inflows		Outflows	
Survivors from 1980-85	56,618	Deaths in 1985-90	3,258
Births 1985-90	3,867	Survivors to 1990-95	57,411
Net immigrants 1985-90	184		
Total	60,669	Total	60,669

Source: modelled on Table III.4 in Stone 1971
statistics from OPCS (1992c), Table 5, p.40. Crown Copyright

Notes:

1. the time intervals all refer to mid-year to mid-year
2. the net immigrants term includes "Net civilian migration" and "other changes"
3. the figures for 1991 in the OPCS (1992c) Table 5 are not, as yet, consistent with previous estimates built forward from the 1981 Census
4. the figures in the table are in 1000's

Early population accounts were defined as tables of population flows. However, not just any table will do. A population account table must include a representation of all possible states that people could have originated from and all possible states that they can exit to. In practice, this means that a "rest of the system of interest" state must be recognised. Usually this "rest of the system" takes the form of "all countries outside the country of interest" although this often leads to complications in a multi-country nation-state such as the UK. A good deal of pioneering work on population modelling failed to recognise the necessity of closing the system of interest properly and the urge to simplify continues (eg Shen 1991). A good test to apply when constructing a population account is to ask whether international movement has been captured or not.

The second feature of closed population accounts is that they help recognise flows between population states (here the term is being used in the sense of defining attribute). They are usually defined in the form of a matrix. Table 5 shows such a matrix and illustrates the general notation to be used. F represents a flow, i represents an initial state of which there are m and a final state of which there are F, so that F^{ij} is the flow of people between initial state i and final state j. In the general case,

Table 5 A matrix of multistate population flows

		Final states				
		1	...	j	...	n
Initial states	1	F^{11}	...	F^{11}	...	F^{in}
	:	:		:		:
	i	F^{i1}	...	F^{ij}	...	F^{in}
	:	:		:		:
	m	F^{m1}	...	F^{mj}	...	F^{mn}

population flows are allowed between all states. If we define the states to be regions, then this is always the case. In other situations there may be restrictions on entry and exit. Persons may enter the married state from the single, never married state, but can never return there.

If the population states are defined as ages, persons can never move from older to younger states. The definition of age is closely linked to that of birth date and time. It is possible to define age-time states so that persons can remain in the same age group at the end of a time interval. However, it is unwise to use such accounts for building any kind of projection model. This is because it can lead to logical inconsistencies such as survival within the age group long after the elapse of time dictates that everyone must have moved out of the age group. You can't get any younger. Both Stone (United Nations 1976) and Rees and Wilson (1977) on occasion adopt this "Peter Pan" approach to demographic modelling.

2.2 Methods of measuring migration

In Table 5 the rather vague term "population flow" has been deliberately used. When actual population account tables are assembled, it is necessary to be more precise about the meaning of these flows. The meaning depends essentially on the way migration between population states has been defined in the instruments used to record migration. Figure 6 illustrates the alternative concepts and their associated methods using graphs of space (reduced to one dimension) against time. The top graph (A) represents the history of six lives lived in two regions over a general time interval (time t to time t+1). The six lives encompass the four possible transitions between "life states": three people exist at the start of the interval and survive at the end; one person exists and then dies; one person is born during the interval and survives at the end; one person is born but also dies in the interval. The three persons in the exist-survive group show three possible behaviours: one stays put in region i throughout, one migrates to region j but then returns to region i and stays there. The third member of this group migrates from region j to region i and stays there to the end of the interval. These alternative behaviours are also possible for the other life-state transitions but are not represented in Figure 6 to keep the diagram simple. The information contained in graph A can only be captured by a full population register in which demographic events are linked to form life histories. Only a few countries such as Sweden keep such records. The Longitudinal Study of OPCS provides some of the information for a 1% sample

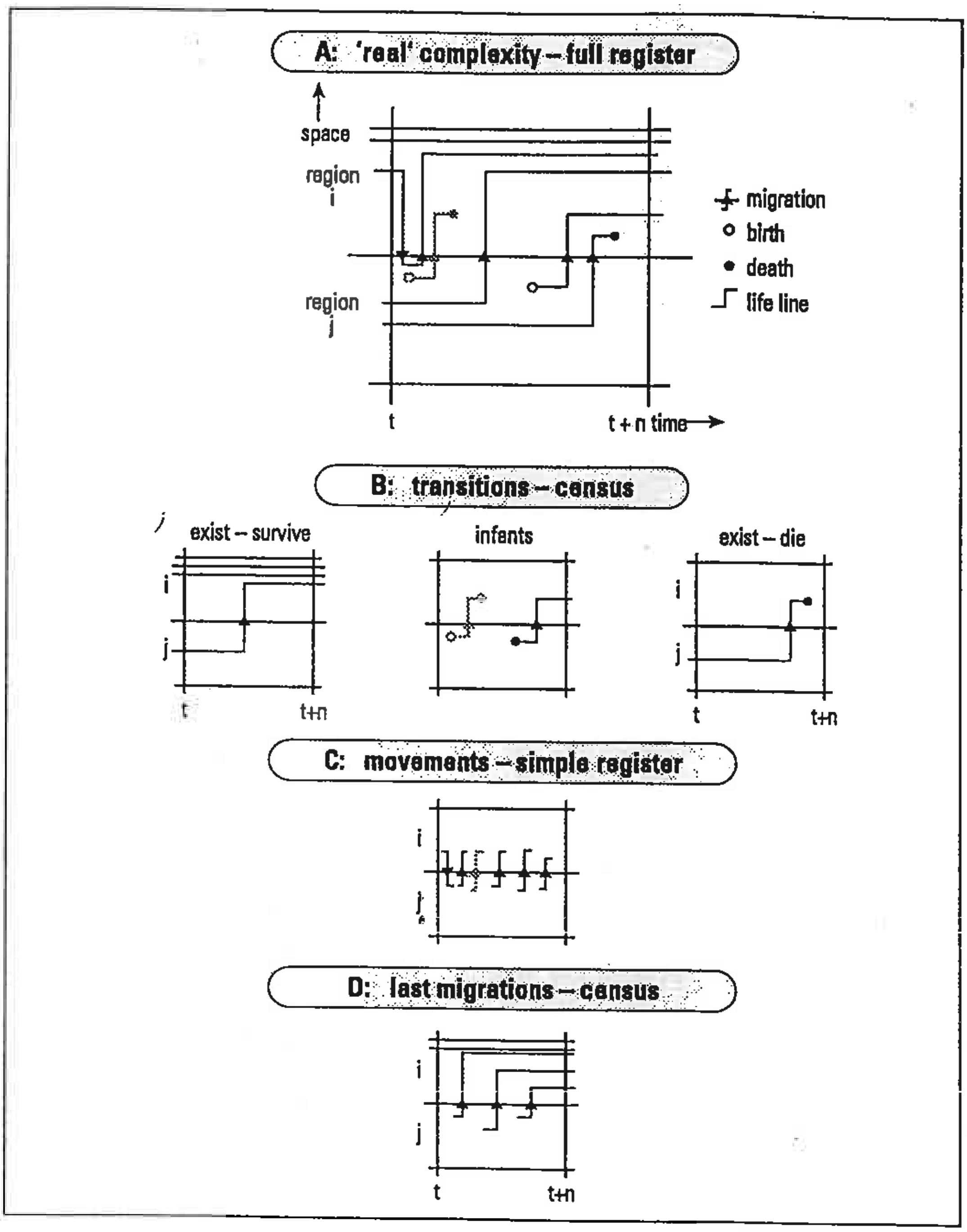

Figure 6: Space–time diagrams illustrating migration concepts

of the England and Wales population but does not currently record migration events.

The commonest method of measuring migration is to ask a retrospective question in a survey or census of the form: "where were you living n years before the date of the Census?", where the interval n is usually one year or five. This makes possible the capture of one migration event out of the six (graph B: exist-survive transitions), with another two persons being classified as stayers in region i. One of these stayers is a return migrant but this migration is not captured normally. There are two infant migrants in our life line sample. By infant in this context is meant persons born during the time interval. In principle it should be possible by asking the question "where were you born?" to ascertain the birth-survive transitions. This has routinely been carried out in United States Censuses since 1940 and more recently in Canadian Censuses since 1971. A birthplace question is asked in UK Censuses and between 1851 and 1951 life time migration tables were routinely published. However, with the introduction of a fixed interval migration question in 1961, the detailed coding of birthplace within each UK country was dropped, and despite lobbying during the consultation periods before the 1981 and 1991 Censuses (Rees 1989) it has not been restored.

Neither of the migrations involving death of the migrant are captured in the Census, nor would it be appropriate to add the relevant question to the death registration form. These are transitions which, in the absence of full population registers, must be estimated rather than counted.

A second method of measuring migration is through systems of registering moves as shown in graph C of Figure 6. These involve the requirement that all citizens must register their change of address with local authorities as happens in the Low Countries and Scandinavia. A change of address can also be imputed from re-registration at a new location for services, as happens with the patient re-registration system of the UK National Health Service (Devis and Mills 1986, Stillwell, Rees and Boden 1992). Such registration systems can capture, in principle, all movements occurring in the time interval, but such movements cannot be linked to the state of the movers at the beginning or end of the period.

A third common method of measuring migration is to ask the question in a census or survey of the form

> "What was your previous usual residence before the current one within the last n years?"

This is a popular question with census takers, particularly in developing countries, because it is easier to answer accurately than the fixed interval question and yields more information on migration. However, the migrants cannot be definitively linked back to their places or origin at the start of the time interval. This characteristic means that such information cannot be directly used in building population accounts and successive authors have argued strongly for use of the retrospective question (United Nations 1970, Courgeau 1980 and Rees 1985).

Population account tables can be constructed using either the transition concept or the movement concept. These are now discussed in turn beginning with the simpler movement account.

2.3 The movement account

Table 6 shows an example of a movement account for a two region UK system in the five year period from mid-1976 to mid-1981. Greater London starts the period with a population of 7 millions; it loses 408 thousand people through death over the five years and nearly 260 thousand emigrations occur to the outside world. Just over 1 million and 50 thousand migrations occur out of the South East. This leaves an accounting residual of 5.3 millions in the South East. To this are added 771 thousand in-migrations from the Rest of

Table 6 An example of a population account based on the movement perspective

State before move in 1976-81	State after move in 1976-81: Internal destinations: Greater London	Rest of the UK	Outside world	Death	Totals
Greater London	5,307,263 (ABM)	1,053,475 (NHSCR)	259,291 (IMS)	408,271 (VS)	7,028,300 (PE)
Rest of the UK	770,541 (NHSCR)	44,395,991 (ABM)	781,210 (IMS)	2,954,158 (VS)	48,901,900 (PE)
Outside World	287,104 (IMS)	603,356 (IMS)	0	0	890,460 (IMS)
Birth	443,546 (VS)	3,136,291 (VS)	0	0	3,579,837 (VS)
Totals	6,808,454 (PE/AC)	49,189,113 (PE/AC)	1,040,501 (IMS)	3,362,429 (VS)	60,400,497 (AC)

Sources: Rees (1984) from OPCS, GRO(S) and DHSS (NI) statistics. Crown Copyright
NHSCR = National Health Service Central Register (UK)
IMS = International Migration Statistics (International Passenger Survey, UK)
VS = Vital statistics
ABM = Account Based Model
PE = Population Estimate
AC = summed from the ACcount

Notes: The time interval is from mid-year (June 30/July 1) 1976 to mid-year 1981 the UK and 287 thousand immigrations. Births in the South East total 444 thousand. When the South East column is summed the result is the South East's population at mid-1981. Note the very high level of emigration from the UK and the relative importance of Greater London in international population exchanges: with 12% of the UK population, Greater London handled 28% of the international migration flows in the 1976-81 period.

A similar set of accounting numbers are given in the table for the Rest of the UK. In the bottom right of the account zeroes are entered. The zeroes in the outside world row are there because the population of the rest of the world that never visits our shores in the period makes no contribution to either the starting or finishing populations of regions of interest. Zeroes are entered in the births row because infants born in the period who die are already counted in the births row and deaths column while infants who emigrate are already counted in the flows to the outside world.

Once the relevant data have been assembled from the sources noted in the table, the items can be entered in the table directly with the exception of the diagonal terms. These must be computed as residuals using the row (or column) relationship. The account table can be used to estimate the final population in the period and is thus a method for updating geographical population counts in the years after a census has been carried out.

It is useful to produce a formal representation of the variables in a movement account and this is done in Table 7. Using this formal representation a variety of projection models have been developed (see Rees 1984, 1989b) though the definition of occurrence-exposure rates for each of the event counts represented in the table. For example, the mortality rate can be defined using the average population in an interval:

$$d^{i} = D^{i}/\tfrac{1}{2}\ (P^{i\cdot} + P^{\cdot i}) \qquad (1)$$

or the migration rate using an origin specific population at risk

$$m^{ij} = M^{ij}/\tfrac{1}{2}\ (P^{i\cdot} + P^{\cdot i}) \qquad (2)$$

it is not appropriate here to expound the associated projection models in detail, but the choice of notation is significant. Different letters have been used to represent the different kinds of events, whereas a common variable is adopted in the transition accounts described below. The superscripts carry the

Table 7 **Algebraic notation for variables in a population account based on the movement perspective**

State before move in interval		State after move in interval: Destination region 1	...	i	...	n	Outside World	Death	Totals
Origin Region	1	R^1	...	M^{1i}	...	M^{1n}	E^1	D^1	$P^{1\cdot}$
	:	:		:		:	:	:	:
	i	M^{i1}	...	R^i	...	M^{in}	E^i	D^i	$P^{i\cdot}$
	:	:		:		:	:	:	:
	n	M^{n1}	...	M^{ni}	...	R^n	E^n	D^n	$P^{n\cdot}$
Outside World		I^1	...	I^i	...	I^n	0	0	$I^\cdot$
Births		B^1	...	B^i	...	B^n	0	0	$B^\cdot$
Totals		$P^{\cdot 1}$	...	$P^{\cdot 2}$	...	$P^{\cdot n}$	$E^\cdot$	$D^\cdot$	$T^\cdot$

Notes: The terms are defined in general as:

R^i = residual term for region i
M^{ij} = migrations from region i to region j
E^i = emigrations from region i
D^i = deaths in region i
I^i = immigrations to region i
B^i = births in region i

$P^{i\cdot}$ = start population in region i
$P^{\cdot i}$ = end population in region i

$E^\cdot$ = total population from all regions
$D^\cdot$ = total deaths in all regions
$I^\cdot$ = total immigrations to all regions
$B^\cdot$ = total births in all regions

$T^\cdot$ = total stocks and flows input/output

0 = term excluded from the account
$\cdot$ = summation over superscript replaced

state information (on locations) for each event count.

2.4 The transition account

We begin an exposition of transition accounts with the formal representation of Table 8, as closely parallel as possible to the movement account representation of Table 7. A common letter is used to indicate a count of persons in all cells. The attributes of the persons are given by the sequence of four superscripts, which have the following general meaning. The first superscript refers to starting life state (existence or birth) in the time interval; the second superscript refers to the associated location; the third superscript identifies the ending life state (survival or death); while the final superscript gives the associated location for the ending life state. Imagine the variables as persons carrying on baggage their backs, the baggage being a portion of their life history (the superscript sequence). The dot superscript indicates there has been a summation of terms using the replaced superscript. For example, the final, surviving population of a typical region i is represented as

$$p^{\cdot\cdot si}=p^{eisi}+\sum_{j\neq i}p^{ejsi}+p^{bisi}+\sum_{j\neq i}p^{bjsi} \tag{3}$$

Table 8

Algebraic notation for variables in a population account based on the transition perspective

Initial state a time interval		Final state in a time interval: Survivors (s), Destination region 1	...	i	...	n	Outside World r	Non-survivors (d)	Totals
Existers Origin region	1	p^{e1s1}	...	p^{e1si}	...	p^{e1sn}	p^{e1sr}	$p^{e1d.}$	$p^{e1..}$
	:	:		:		:	:	:	:
	i	p^{eis1}	...	p^{eisi}	...	p^{eisn}	p^{eisr}	$p^{eid.}$	$p^{ei..}$
	:	:		:		:	:	:	:
	n	p^{ens1}	...	p^{ensi}	...	p^{ensn}	p^{ensr}	$p^{end.}$	$p^{en..}$
Outside World	r	p^{ers1}	...	p^{ersi}	...	p^{ersn}	0	$p^{erd.}$	$p^{er..}$
Birth (b)		$p^{b.s1}$	...	$p^{b.si}$	...	$p^{b.sn}$	$p^{b.sr}$	$p^{b.d.}$	$p^{b...}$
Totals		$p^{..s1}$	...	$p^{..si}$	...	$p^{..sn}$	$p^{..sr}$	$p^{..d.}$	$p^{....}$

Notes on variables:

P =population/persons

p^{eisi}= persons existing in region i, surviving in region i (surviving stayers)
p^{eisj}= persons existing in region i, surviving in region (surviving migrants)
p^{eisr}= persons existing in region i, surviving in region r (surviving emigrants)
$p^{eid.}$= persons existing in region i, who die anywhere (non-survivors)
$p^{ei..}$= persons existing in region i (start population)

p^{ersi}= persons existing in region r, surviving in region i (surviving immigrants)
$p^{erd.}$= persons existing in region r, dying in our country (non-surviving immigrants)
$p^{er..}$= persons existing in region r, who immigrate (total immigrants)

$p^{b.si}$= persons born anywhere, surviving in region i (infant survivors)
$p^{b.sr}$= persons born in our country, surviving in region r (infant surviving emigrants)
$p^{b.d.}$= persons born, who die (infant non survivors)
$p^{b...}$= persons born (total infants)

$p^{..si}$= persons surviving in region i (final population)
$p^{..sr}$= persons surviving in region r (total surviving emigrants)
$p^{..d.}$= persons dying (total non-survivors)

That is, the surviving population in a region is the sum of surviving stayers, all surviving in-migrants (including immigrants), surviving infant stayers and all surviving infant in-migrants (including immigrants).

From the elements of a transition account it is relatively easy to compute transition probabilities for use in projection models. For example, the probability of that a person in existence in region i at the start of the period will survive in region j at the end is given by:

$$h^{eisj} = p^{eisj}/p^{ei..} \tag{4}$$

where h is the letter representing transition probability. These are proper probabilities because, for an account table only, they necessarily sum to unity across the rows of the table:

$$\sum_j h^{eisj} + h^{eid.} = 1 \tag{5}$$

that is, as long as all possible ending states have been specified.

How easy is it to find and enter the data for Table 8's transition account? End of period census populations and census migrants (internal to a country and immigrants) slot in easily as final population terms, $P^{\cdot\cdot si}$, and surviving migrant terms P^{eisj}. Some ingenuity must be exercised to estimate the surviving emigrant terms, P^{eisr}, from emigration counts through application of a survival fraction. The infant survivors, $P^{b\cdot si}$, can be extracted from end-of-period census populations in the youngest age group. Surviving stayers, P^{eisi}, if not directly available from the census migration tables, can be computed by subtracting all in-migrants from the final population in the account.

We are now left with the column of non-survivors and the starting totals in the right most column, the first block of which are initial populations. If these are known from a start of period census (rather rarely) or from a population estimate series, then non-survivors can be worked out as a residual. Residuals, however, tend to cumulate errors in all the other terms, and it would be preferable to seek direct information about deaths to regional populations.

Here a subtle difficulty is faced: there is plenty of detailed and reliable information on regional mortality but it is not quite what is needed in the transition account of Table 8. What is published are counts of deaths by region of death, that is variables of the form $P^{\cdot\cdot di}$ or its components $P^{e.di}$ and $P^{b.di}$ (non-infant and infant deaths), and not counts of deaths by the prior location of the deceased, the variable $P^{eid.}$, irrespective of the location at time of death. The count of deaths $P^{\cdot\cdot di}$ is an event count which we labelled D^i in the movement account (Table 7).

A variety of different methods have been developed to overcome this difficulty, including the accounting models of Rees and Wilson (1977), the multistate models of Ledent (1980) and the recent accounting models of Shen (1991) and Shen and Spence (1993). The key step is to expand the representation of non-survivors into a matrix paralleling that of survivors, as is done in Table 9. The row sums of this matrix are the total non-survivors needed in the earlier transition account and the column sums are the observed regional mortality totals. The accounting models aim to estimate all the interior terms using occurrence-exposure death rates based on the observed death counts. Such accounting models contain elements of both transition and movement methodology. As in so many fields, these hybrid models have provided vigorous stimuli to conceptual advances and preceded the clear distinctions between movement and transition perspectives which have been made here.

Table 9

Algebraic notation for variables in a population account based on the transition perspective: expanded form for non-survivors

Initial state in a time interval		Final state in a time interval: Non-survivors — Destination region 1	...	i	...	n	Outside world r	Sub totals	Totals
Existence (e)									
Origin regions	1	p^{e1d1}	...	p^{e1di}	...	p^{e1dn}	p^{e1dr}	$p^{e1d.}$	$p^{e1..}$
	:	:		:		:	:	:	:
	i	p^{eid1}	...	p^{eidi}	...	p^{eidn}	p^{eidr}	$p^{eid.}$	$p^{ei..}$
	:	:		:		:	:	:	:
	n	p^{end1}	...	p^{endi}	...	p^{endn}	p^{endr}	$p^{end.}$	$p^{en..}$
Outside world	r	p^{erd1}	...	p^{erdi}	...	p^{erdn}	0	$p^{erd.}$	$p^{er..}$
Birth (b)									
Origin regions	1	p^{b1d1}	...	p^{b1di}	...	p^{b1dn}	p^{b1dr}	$p^{b1d.}$	$p^{b1..}$
	:	:		:		:	:	:	:
	i	p^{bid1}	...	p^{bidi}	...	p^{bidn}	p^{bidr}	$p^{bid.}$	$p^{bi..}$
	:	:		:		:	:	:	
	n	p^{bnd1}	...	p^{bndi}	...	p^{bndn}	p^{bndr}	$p^{bnd.}$	$p^{bn..}$
Outside world	r	p^{brd1}	...	p^{brdi}	...	p^{brdn}	0	$p^{brd.}$	$p^{br..}$
Sub totals		$p^{b.d1}$	...	$p^{..di}$	...	$p^{..dn}$	$p^{..dr}$	$p^{..d.}$	$p^{b...}$
Totals		$p^{..d1}$	...	$p^{..di}$	...	$p^{..dn}$	$p^{..dr}$	$p^{..d.}$	$p^{....}$

Notes on variables:

p^{eidi} = persons existing in region i, dying in region i (non-surviving stayers)
p^{eidj} = persons existing in region i, dying in region j (non-surviving migrants)
p^{bidi} = persons born in region i, dying in region i (non-surviving infant stayers)
p^{bidj} = persons born in region i, dying in region j (non-surviving infant migrants)

Table 10 illustrates an estimated transition account for another British spatial system and indicates the derivation of each of the terms. Full details of the methodology of accounts based modelling and an associated program are provided in Rees (1981). The software is designed to provide the user with a wide variety of choices for filling in population accounts so that virtually all data situations are covered. A later program (Rees 1984) does the same for movement accounts.

2.5 Age disaggregation

For practical applications in population projection, population accounts must be disaggregated by age and sex. In fact, this simplifies matters considerably because we need to adopt only the non-birth part of the accounts for non-infant ages; infants can be handled in slightly adapted versions of the non-birth part of the accounts with regional births replacing initial populations for example. An additional model for generating births must, however, be added.

However, before using any age-disaggregated data as input to an account or to a projection, the user must be absolutely clear as to the age-time scheme used for recording the data and the age-time scheme appropriate for the population model being used, which will always

Table 10 An example of a population account based on the transition perspective

Initial state in a time interval	Survivors at Census 1971			Non-survivors 1966-71			Totals
	South East	Rest of GB	Outside World	South East	Rest of GB	Outside world	
Existence at Census 1966							
South East	14,461.1 (ABM)	653.1 (MT)	648.3 (IMS)	921.0 (ABM)	11.7 (ABM)	3.4 (ABM)	16,698.8 (PC)
Rest of GB	633.3 (MT)	32,791.6 (ABM)	728.5 (IMS)	7.2 (ABM)	2,110.5 (ABM)	4.3 (ABM)	36,275.4 (PC)
Outside World	551.4 (MT)	534.5 (MT)	0	2.9 (ABM)	3.4 (ABM)	0	1,092.1 (AC)
Birth in 1966-71							
South East	1,280.4 (ABM)	38.3 (EMT)	42.2 (EIMS)	23.4 (ABM)	0.3 (ABM)	0.4 (ABM)	1,385.0 (VS)
Rest of GB	35.6 (EMT)	2,997.3 (ABM)	48.2 (EIMS)	0.3 (ABM)	56.9 (ABM)	0.4 (ABM)	3,138.8 (VS)
Outside World	26.4 (EMT)	36.1 (EMT)	0	0.2 (ABM)	0.3 (ABM)	0	63.1 (AC)
Totals	16,988.1 (PC/AC)	37,051.0 (PC/AC)	1,467.2 (AC)	955.1 (VS)	2,183.3 (VS)	8.5 (AC)	58,653.2 (AC)

Sources:

Rees (1977, Table 4) from OPCS and GRO(S) statistics. Crown copyright

MT = Migration Tables from the 1971 Census
IMS = International Migration Statistics
EMT = estimated from MT
EIMS = estimated from IMS
ABM = estimated using account based model
VS = Vital Statistics
PC = Population Census
AC = summed from the ACcount

Notes:

1. The account is adjusted to 1971 Census populations
2. GB = Great Britain
3. The populations are in 1000s

be the "period-cohort" scheme in the case of projection.

Age-time diagrams have a long history in demography are conventionally known as "Lexis diagrams", after their German inventor (Lexis 1875). However, recently Vanderschrick (1992) has asked us to revise our views about the origin of the form of the diagram currently in common use (see Figure 7). Age-time diagrams of this form were actually introduced by the French demographer Pressat (1961), since the original versions by Lexis and contemporaries used year of birth rather than age as the vertical axis of the graph. In addition, Vanderschrick points out that, in the year in which Lexis was writing his monograph, a proto-form of the diagram had already been produced by the German demographer, Becker (1874). Contemporaneously a Dutchman, Verweij (1874 and Verwey 1875) was writing his doctoral thesis in which a diagram very similar to Figure 7 appeared (Vanderschrick 1992, p.1253). Vanderschrick reaches no firm conclusion as to the provenance of the Becker-Lexis-Verweij-Pressat diagram, but he reminds that we need always to look behind the simple labels on new inventions; most are the product of interacting minds rather than the inspiration of an isolated individual.

So what does our age-time diagram (Figure 7) show? It

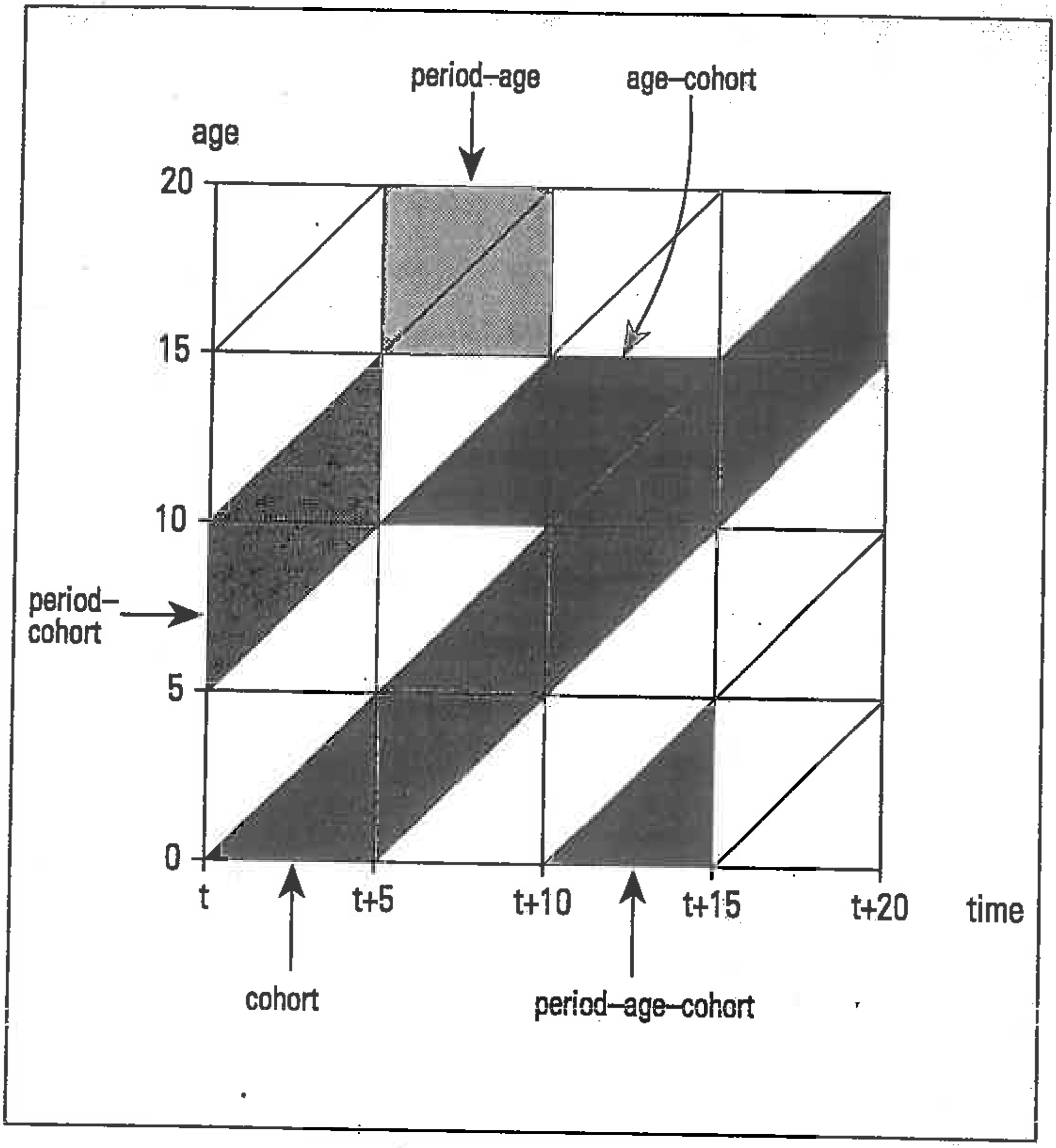

Figure 7: An age–time diagram showing different demographic observation plans

displays the different age-time areas that can be used to record demographic movements or transitions. If these are recorded by a double classification of date of birth and age at time of event, then they can be placed in the period-age-cohort triangle on the diagram. Events counted in this way can be reassembled into period-cohort counts or age-cohort counts or into cohort sequences. Unfortunately, many demographic offices, including those in the UK, still record events in period-age form, and estimating assumptions are needed to reassemble the data in one of the other observation plans, which are more appropriate for demographic analysis.

3. POPULATION PROJECTION

Population accounts provide us with a framework for measuring demographic change in the past, but they also help determine the choice of population projection method. In this section of the paper some general observations about the range of methods available. Exact details of the methods will be found elsewhere.

Projection involves the calculation of future population sizes and attributes based on assumptions about future trends in fertility, mortality and migration. The term forecast is usually reserved for projections invested with some judgment that they are most likely. It is important to be able to make prognoses of the future population in order to be able to plan for future developments. This is essential in pensions planning, in insurance calculations, in health and educational planning, in infrastructure investment and so on. Decisions being taken now about pension contributions, insurance subscriptions, teacher education, housing plans, transport projects all need to take a view, over varying time horizons, of the likely future population, recognising, of course, that a great many non-demographic factors will also be important.

3.1 The ingredients for projection

A projection involves the following ingredients:

(1) a model of the population system and its dynamics, (2) software to implement the model, (3) a benchmark database for some period in the recent past (a population account if available), (4) methods for developing scenarios (assumptions about how the component inputs are likely to change) and (5) the scenarios or assumptions themselves. All these are brought together to produce projections of the future population.

Figure 8 shows the ingredients of an integrated population projection system for local authority wards developed for the City of Swansea in 1990 by GMAP Ltd and the School of Geography (Rees and Rees 1990; Rees, 1991a). The projection model is built from the movement perspective and is implemented in Fortran code on a personal computer. The benchmark data are assembled for the years 1981-86 in full though population estimates were prepared to 1989. Note that the benchmark database is conceived of in two parts: the raw database (the demographic statistics that are available from official agencies) and the estimated database which contains population rates and parameters suitable for entry into the projection model. The two parts are connected by estimation models (a subject taken up in section 6). The key feature of the projected database was the estimation of the size of the population of electoral age for

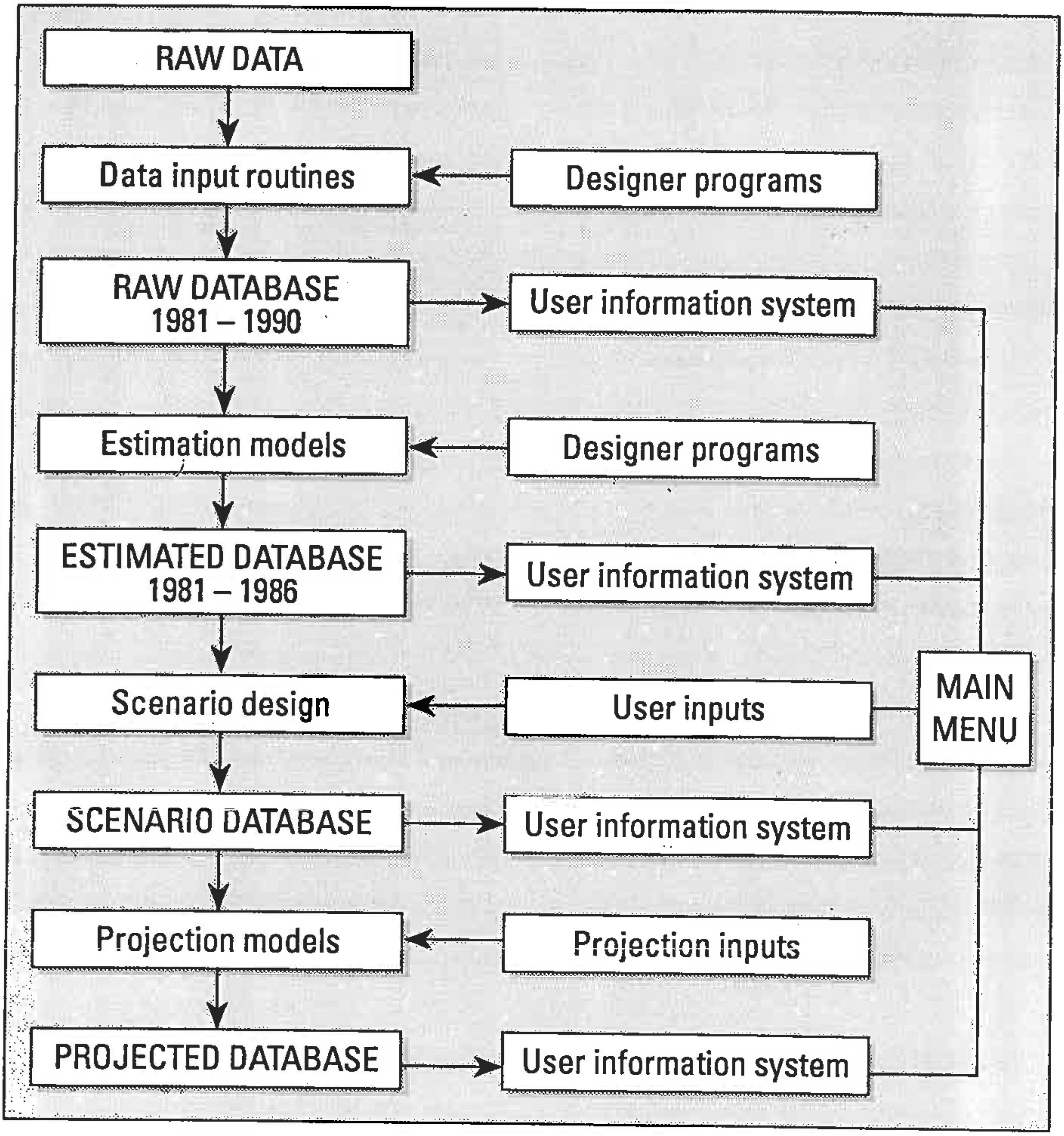

Figure 8: The structure of the Swansea model

each ward as input to negotiations being conducted between the City of Swansea and the Boundary Commission for Wales.

3.2 The form of the projection model

3.2.1 Accounting perspective: rates or probabilities

The structure of the model used for projecting the population will be based on which, of two accounting perspectives, movement or transition are explicitly or implicitly adopted. If the movement perspective is adopted, the model will use occurrence-exposure rates of mortality, fertility and migration applied to populations at risk based most conveniently on the average population in a time interval. If the transition perspective is adopted, the model will use transition probabilities of staying and survival, migration and survival, and of non-survival, together occurrence-exposure fertility rates.

3.2.2 Number of interacting populations

A key decision in constructing a population projection models is to determine how many populations will be allowed to interact. Should the projection model deal with a single region, with a region and the rest of the country, or with a set of regions? Single region models normally handle migration in net form. Biregional and multiregional models represent migration in gross form, projecting both the flow from region i to region j and the reverse flow from j to i. Rogers (from 1968 to 1991) has argued strongly for the logic of the gross flow and multiregional approach, and his arguments are now widely accepted.

3.2.3 Parameterisation of age schedules

When a multiregional approach to projection is adopted along with a fine level of age disaggregation, very large arrays of variables can result which cannot be reliably measured or estimated. For the example, in a projection model constructed for Bradford wards (Rees 1991d, 1992b) adoption of a full multiregional approach would have resulted in the need to estimate transition probabilities for 30 origin wards by 30 destination wards for 91 ages and two sexes: some 163,800 variables would have been involved. Instead schedules of demographic rates by age can be replaced by the parameters of a function that describes the variation. This approach is known as "parameterisation" and has been pioneered by Rogers and his fellow workers (Rogers and Castro 1981a, 1981b; Rogers and Planck 1983; Rogers 1986). Most widely used is the Rogers-Castro model of migration rates by age which decomposes the schedule into four components and a constant, and represents the components by single or double exponentials.

3.2.4 Aggregation and decomposition of projection models

In earlier work Rogers (1976) explores the savings in terms of inputs that would result from either partitioning or decomposing the multiregional projection model and the losses in terms of accuracy compared with the full model. One useful decomposition was to model regional shares using an aggregate multiregional model but age composition and total size using single region methods. In a recent projection of the populations of the 71 regions (NUTS level 1) of the European Community (EC), Rees, Stillwell and Convey (1992) adopt a hierarchical approach Figure 9. Migration between the 12 EC member states is modelled using a multiregional, all age model. Total in-migration to a member state from all others is then distributed across the regions of the member state, and total out-migration is treated similarly. These migration components are combined with the results of a multiregional model applied to the regions within end member state and a distribution of the member state's assured net extra-Community migration. Age disaggregation is effected using model migration schedules.

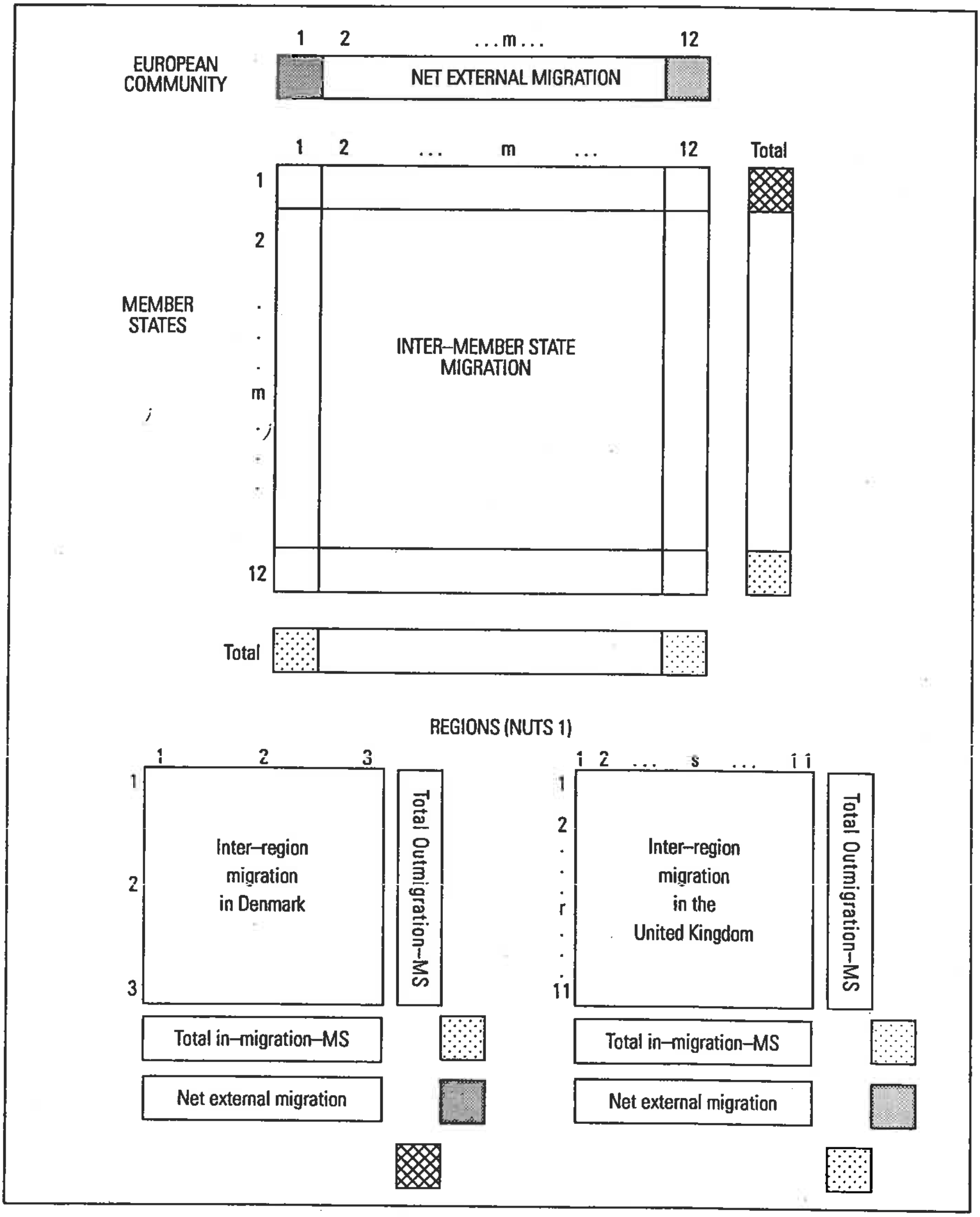

Figure 9: The hierarchical structure of a projection model for 71 European Community regions

3.2.5 Parameterisation of the interactions

In the Bradford model, part of a wider project for West Yorkshire local authorities being carried out by GMAP Ltd and the School of Geography, both hierarchical decomposition and parameterisation are adapted to solve the "too many variables" problem. Migration flows between wards are projected on an all age basis and only broken down by age when summed to yield in-and-out migration totals (using estimated migration probability schedules rather than model generated schedules). In addition, parameterisation is taken one step further. Instead of projecting the migration between wards as the product of a migration probability and an origin ward population, a set of spatial interaction models is used that connect some of the migration flow to developments in the housing market, specifically new housing and housing demolition as illustrated in Figure 10. This makes the outcomes of the projection sensitive to policy assumptions, as well as reducing the number of inputs to the migration part of the projection.

3.2.6 Solution methods

One problem that occurs in projection models derived from the movement perspective and from the extended hybrid transition perspective is that the population

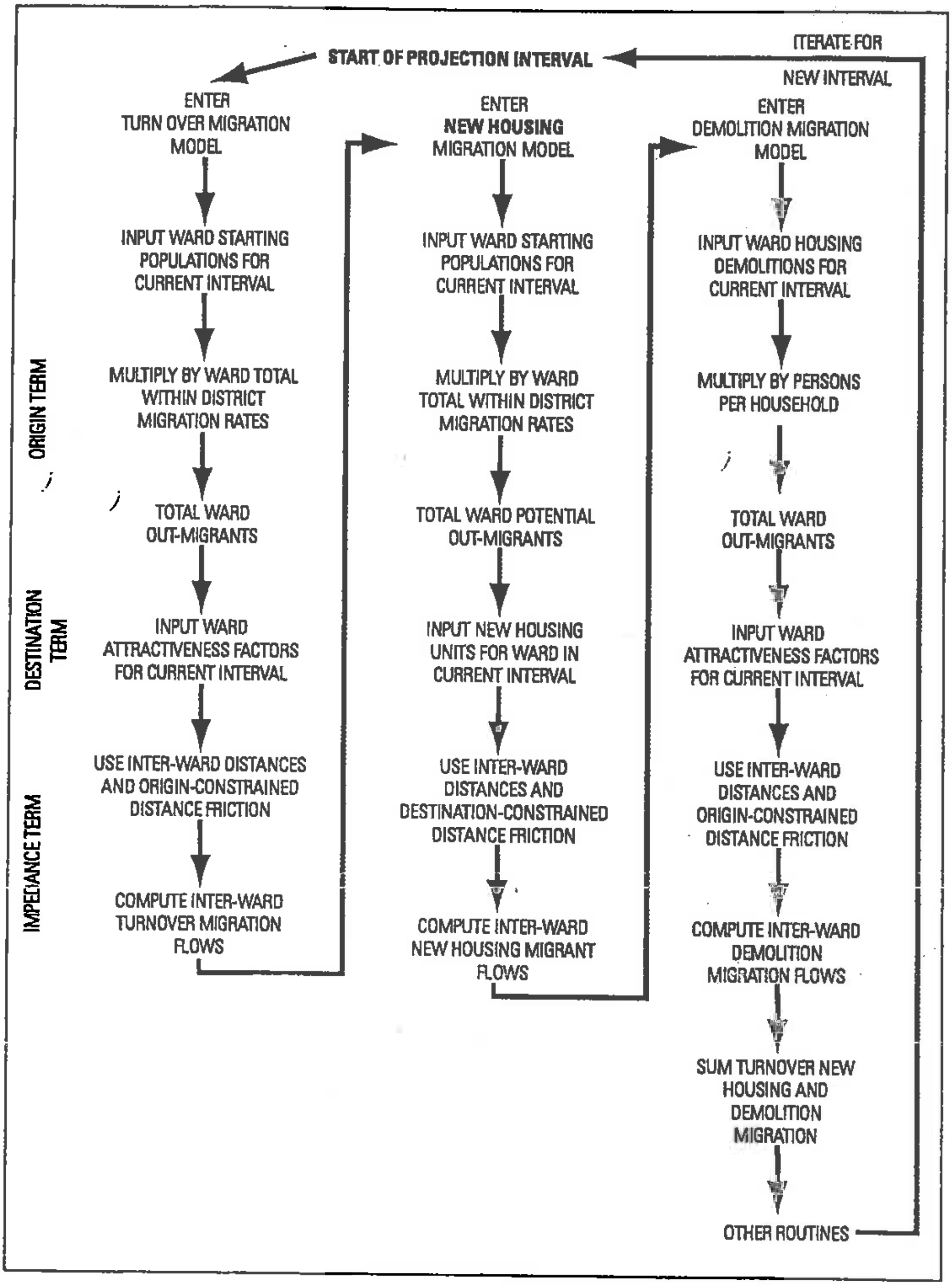

Figure 10: The spatial interaction models used to project inter-ward migration within West Yorkshire districts

at risk of experiencing events during the projection interval is unknown. This problem does not occur in the pure transition perspective, when the start populations of origin regions are the appropriate populations at risk.

There are two solutions. The most favoured is to solve the projection equations so that a one step model is derived. Willekens and Rogers (1978) and Willekens and Drewe (1984) contain elegant derivations of these and a further explanation is given in Rees (1989b). A less favoured solution is to use iteration which is mathematically untidy but has the advantage making possible the construction of very flexible software that is, in effect, a model-building kit for the user rather than a fixed, prescribed model (Rees 1984). In practice, iteration is not avoided in the one-step models, it is merely hidden in the procedures for computing matrix inversions.

3.2.7 Closure methods

A final set of choices which the projection model builder must make is to determine the method of system closure: that is, how interactions with the rest of the world should be handled (Rees 1986a). Six alternatives are listed in Table 11. Methods (b) and (c) are popular in national, single population models,

Table 11 **Alternative methods of system closure for subnational population projection methods**

Closure method	Comments	Applicability to projection methods
(a) World Model	The rest of the world is treated as a subnational unit	Multi-unit model
(b) Net immigration rates/gross emigration and immigration rates	Rates are applied to the subnational unit population at risk only	All methods
(c) Net migration flows/ gross emigration and immigration flows	Exogeneous projection of the the external migration flows needed	All methods
(d) Emigration rates and immigration flows	Reflects normal degree of political control over respective migration streams	All methods
(e) Emigration flows and immigration rates	Probably unrealistic; not recommended	All methods
(f) Ignore external migration	Unrealistic: use only in the absence of any data	All methods

while option (d) is used frequently with multiregional projection models. Alternative (a), in which the population of the rest of the world is treated as just another regions, is easy to implement given a handy copy of the United Nation's latest **World Population Prospects**. However, it is really only useful to explore a world where national frontiers do not serve as serve barriers to international migration.

3.3. Scenario development

The final aspect of population projection that needs to be briefly discussed is scenario development. Here we concentrate on technical aspects and discuss the generation of assumptions in section 4.

To project the population normally requires that trajectories of the input rates and parameters be specified beforehand. The simplest way would be to develop trajectories of all the individual rates and probabilities for all regions, inter-region flows, ages and sexes. This is a formidable task which is usually avoided. Instead the problem is reduced to one of specifying the way in which a few summary parameters may change. For example, instead of specifying how 30 age specific fertility rates (AS FRs) for the 30 wards of Bradford will change over the next 15 years, the user of the projection model would specify the trajectory of the total fertility rates (TFRs) for the wards instead, assuming constancy of the relative distribution of TFR across the ASFRs. If there was only a need to look quickly at the outcomes of high or low fertility scenarios, the same change scenario could be applied to all wards. The minimum input for scenario development in this case would be to use one percentage change per year factor across all wards and all time intervals.

4. APPLICATIONS: MORTALITY AND MIGRATION TRENDS

The assumptions that drive the outcomes of well specified projection models must be derived from careful study of recent trends and analysis of the factors explaining those trends. This is the main focus of most contemporary work in population geography and demography. Two brief examples are reviewed here which serve to emphasise how vital it is to have up-to-date knowledge of trends and to understand what surprises might be lurking therein.

4.1 Trends in national mortality

Life expectancy has improved steadily in the United Kingdom this century although most other western developed countries have had a better performance in that respect (Rees 1989a). Figure 11 documents experience in the constituent countries of the UK since 1961. About four years have been added to life expectancy over that period. The experience of the period after 1976 has been more favourable than the period before 1976. The gain in life expectancy years per calendar year for UK males was 0.11 in the 1961 to 1976 period but 0.24 of a year per year thereafter, while the corresponding statistics for females were 0.09 and 0.24 of a year per year. One might say that the UK population in the post 1976 period was a quarter of the way towards immortality (represented by achievement of

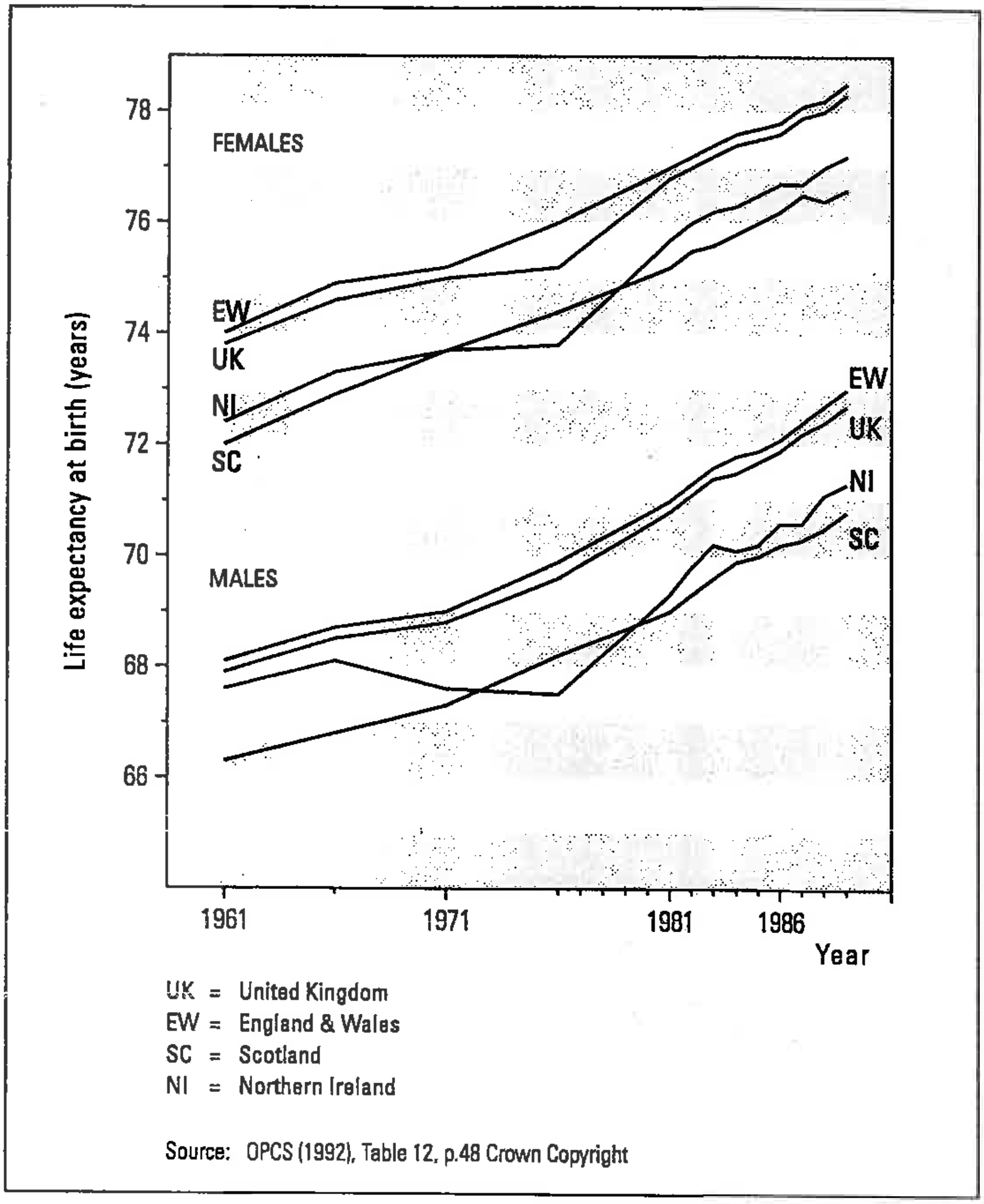

Figure 11: Trends in life expectancy at birth 1961 – 89, UK

gains of one year of life per calendar year elapsed).

Would we be right in assuming a continuation of that gain when carrying out population projections? A detailed analysis by cause of death would be needed for a proper judgment, but even if the gains continued for another 25 years, the UK population would not have caught up with contemporary Icelandic, Japanese and Swedish populations.

There are, however, one or two nasty surprises lurking in the cause of death statistics which need to be taken into account. The spread of the HIV/AIDS epidemic is one factor which is increasing mortality rates, particularly in the 20 to 50 age range. Table 12 reports on one national projection which shows the effect of assuming an epidemic reaching an equilibrium of 30,600 AIDS cases per year by the year 2000. By 2051 this means a cumulative loss to the population of 2 million persons. Note that if all demographic rates remain constant at their 1981-86 level the population will decline significantly in the twenty first century. A second source of increased deaths are the increased deaths from lung and associated cancers as the cohort of women smokers reaches the vulnerable years. More recently a much more sophisticated model for simulating the HIV/AIDS epidemic in the UK has

Table 12 Projected populations for the UK, 1981-2051 under four scenarios

Scenario	Year			
	1981	2001	2026	2051
A Constant rates	56,162	57,338	55,269	49,335
B Trended rates	56,162	59,190	63,391	66,760
C Latest migration rates	56,162	59,896	65,562	70,664
D Plus AIDS rates	56,162	59,656	64,458	68,650

Source: Rees (1991), Table 14

Notes: Populations are in 1000s

A Constant rates at 1981-86 level

B Fertility trended to replacement. Mortality improvement steady

C Scenario B but with higher net immigration reflecting mid-80s rise

D Scenario C but with added mortality rates due to AIDS deaths been constructed by Williams (1992, 1993). This model integrates a multi-Health Region demographic model utilising the movement perspective with an epidemiological model that incorporates all of the significant transmission routes and simulates inter-regional spread. Figure 12 summarises the results of the baseline model and for a model in which all assumptions of behavioural change (fewer sexual partners, safer sex, use of sterile needles) have been removed. The incidences (number of new cases) of the both HIV and AIDS are projected from a benchmark in 1987 over the period 1988 to 2000 for six gender/behaviour groups. The model projections for 1988-91 show a very close fit to observed data (Williams 1993, Table 34, p.60).

The Figure 12 graphs are ordered in terms of the three main sub-epidemics: among homosexuals (the first in time), among intravenous drug users (the second) and among heterosexuals (the third). The baseline model, incorporating the best estimates from the HIV/AIDS literature of the key infectivity and behavioural parameters, shows that the male homosexual epidemic of AIDS will peak in the mid-1990s and will then follow with a lag of 10-11 years (the mean incubation time) the declining curve of HIV incidence. The epidemic among male and female intravenous drug users exhibits the same decreasing incidence as among gays, but this

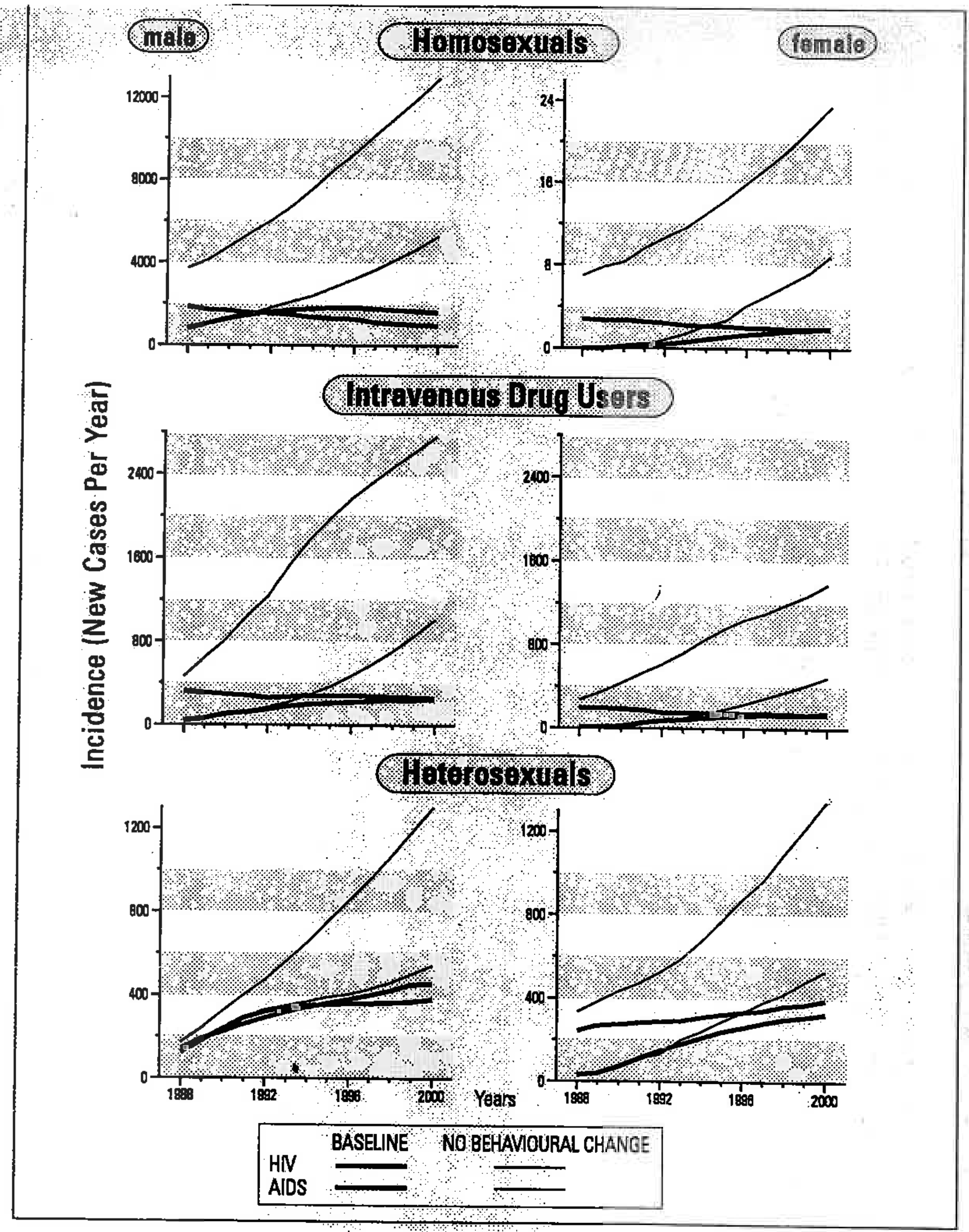

Figure 12: The projected incidence of HIV and AIDS in the UK by behavioural groups, 1988 – 2000: baseline model and model with no behavioural change (after Williams 1993)

is due mainly to safer injecting. Heterosexual spread within the IDU community becomes more important. Currently, the epidemic among heterosexuals continues to augment - the slowing down of the AIDS curve for males is connected with the timing of the epidemic in the haemophiliac group (with virtually no new infections after 1985).

The second set of curves show what might happen if current behavioural restraints were relaxed (more partners, unsafe sex, more needle sharing). The impact is startling. By 2000 this model scenario forecasts 165 thousand infections in UK compared with 55 thousand from the baseline model.

4.2 Spatial differences in mortality

Figure 11 showed no significant convergence in the mortality experience of the constituent parts of the UK, and a more detailed study for the 1980-86 at county/Scottish region scale has confirmed this (Rees 1989). The spatial variation in mortality experience across the country is shown in the four maps of female mortality presented in Figure 13.

The top left hand map shows life expectancy at birth while the other maps show the standardised mortality ratios (SMRs), normalised to a UK standard (not

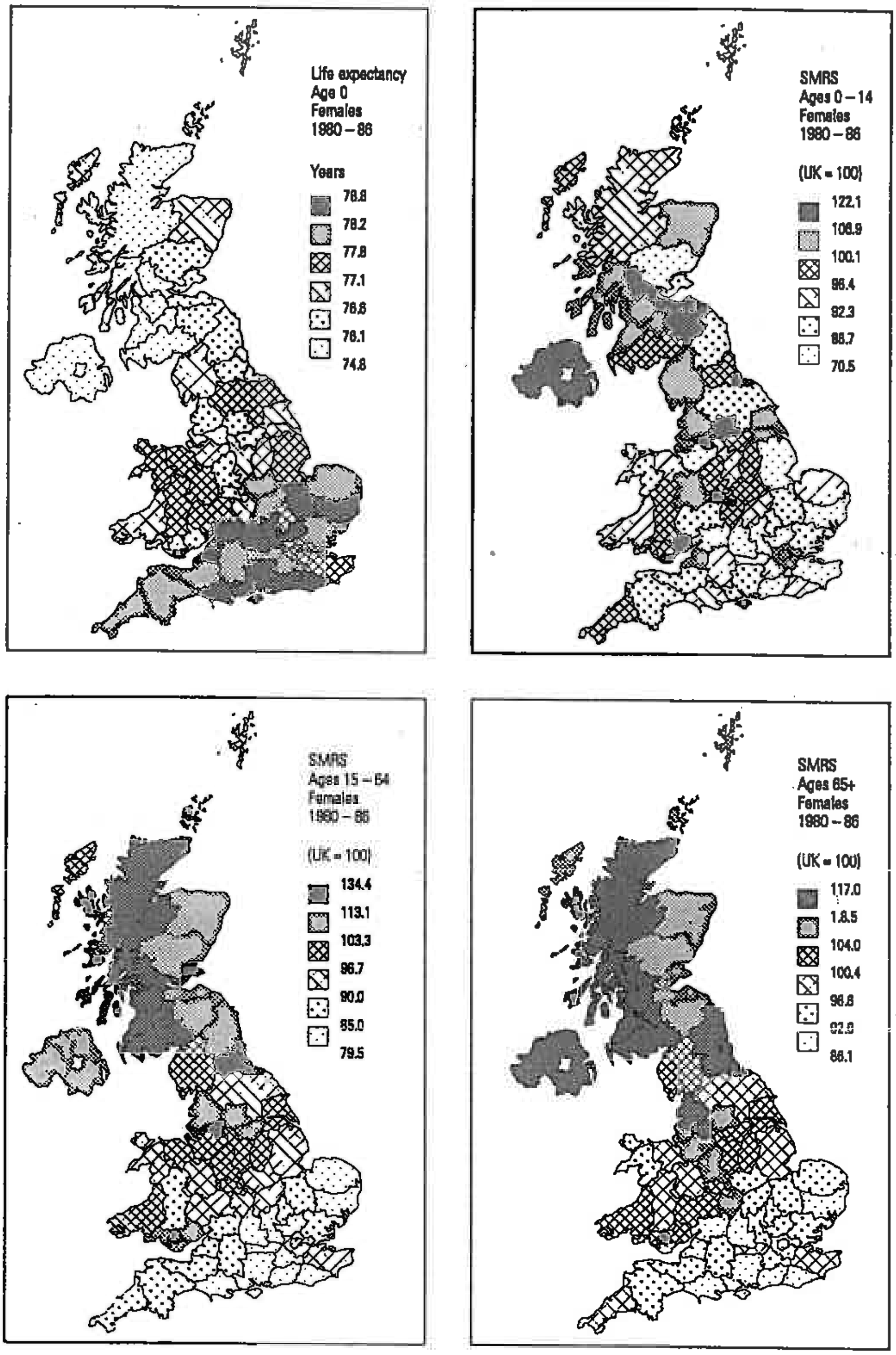

Figure 13: Life expectancy at birth and three standardized mortality ratios, UK, females, 1980 – 86

normally provided in official computations). The shading schemes are arranged so that areas with worst mortality regions are shaded more darkly on both the life expectancy and SMR maps.

Three of the maps show a clear and consistent set of patterns: (1) a gradient from low mortality in the South and East of the country rising to high mortality in the North and West, (2) an urban-rural contrast in most regions with higher mortality in the denser counties and lower mortality in the less dense. The areas which fare worst are most of Scotland, Northern Ireland, the metropolitan counties of Northern England and the South Wales valleys. These mortality maps are closely associated with the pattern of socio-economic deprivation across the country.

These patterns are less clear cut on the map of SMRs for 0-14 year olds. At these ages it is likely that good health care can overcome some of the disadvantages of socioeconomic deprivation: the position of the Northern region counties, Tayside and Fife is much better on this map than on the others.

4.3. Migration and regional economic development

Most projection models for subnational areas make very simple assumptions of constancy of migration pattern.

The official subnational projections for England assume continuation of the pattern of inter-area transfers observed in the year prior to the 1981 Census updating only for the volume of migration for changes in the 1980s (Boden, Stillwell and Rees, 1991). Similarly, the projections used by the Regional Policy Directorate of the European Commission (CEC 1991, Haverkate and Van Haselen 1990) assume constancy over time for the migration pattern. The Statistical Office of the European Communities is currently seeking to improve on these assumptions in a major project on "Regional Population and Labour Force Scenarios" (November 1992-May 1993), while the Office of Population Censuses and Surveys will revise the English sub-national projections as soon as the Special Migration Statistics from the 1991 Census are available.

Figure 14 shows why such revisions are needed. It tracks for the period since mid-1975 the net migration between broad regional divisions of the country. Very large fluctuations have occurred in the net balances. Although most of the post-1975 period has seen an average loss through migration between the North and the South of 50 thousand a year, since 1989 this has disappeared.

What appears to lie behind the fluctuations (though

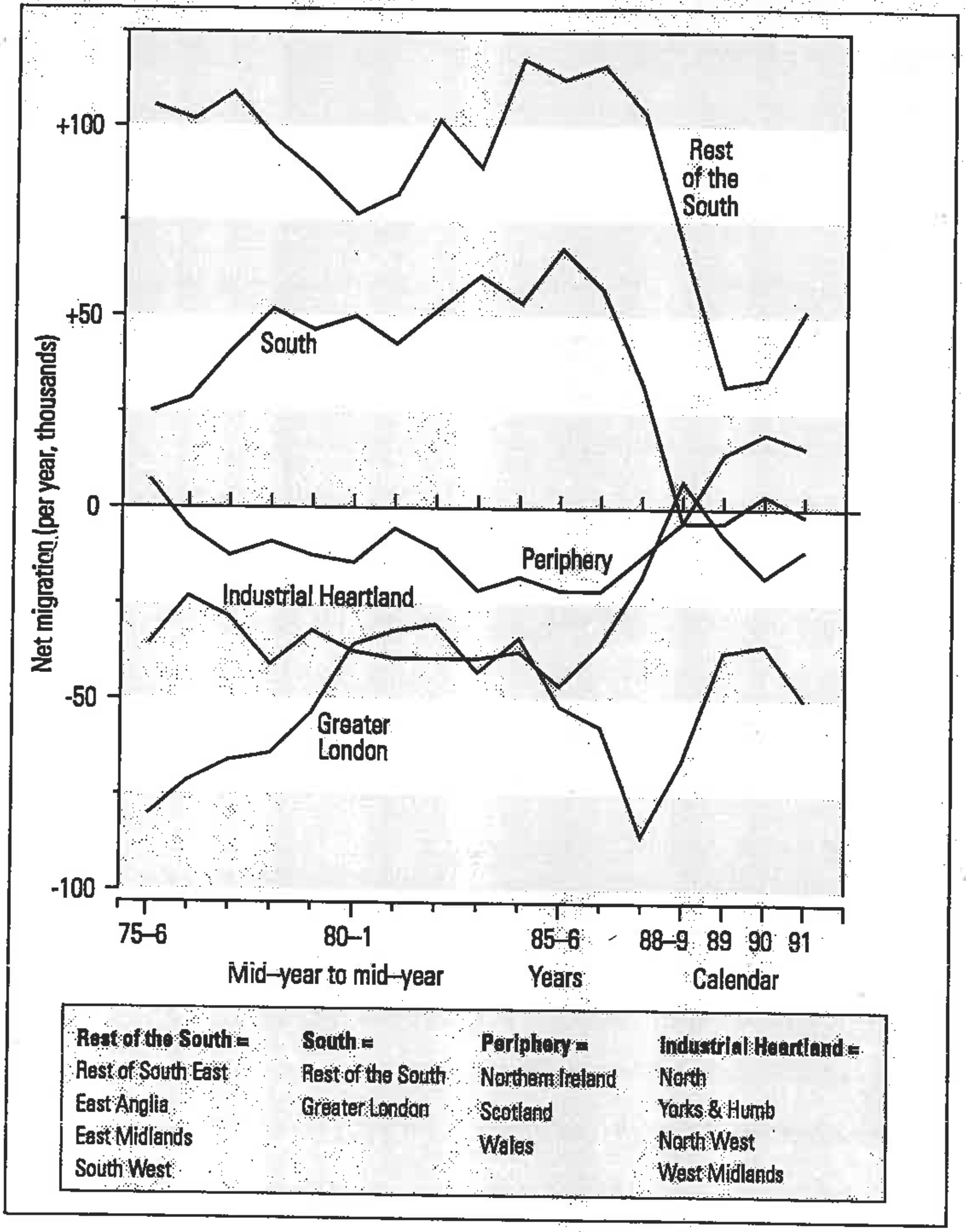

Figure 14: Net internal migration for broad regional divisions, 1975 – 91

this is based on casual observation rather than sustained analysis) is the differential timing of the entry of regions into economic booms or busts coupled with secular structural adjustments. The recession of the early eighties saw the decline in employment in manufacturing begin and bite hardest in the North; the Southern regions were first out of the recession and into the middle and late eighties boom. The early 1990s recession has, however, been led by decline in the service industries of the South East first, with the recession spread northwards (Green et al 1993; Tickell, 1993).

However, net migration is not just a response to the business cycle in a region but also to the relationship between regions. Compton (1992), for example, shows that net migration out of Northern Ireland has been highest in times of relative prosperity in the province because of the availability of more attractive job opportunities on the British mainland.

These relationships between regional economic development and migration need to be incorporated into projection models in the United Kingdom.

5. APPLICATIONS: EVOLUTION OF THE ELDERLY

While the previous section of the paper has focused on trends and patterns for geographically defined populations, it is also valuable to concentrate attention on populations defined using non-geographical attributes. Populations defined in terms of age are particularly interesting, because most of us will experience membership of the different groups over our lifetimes (God willing): there is no way of reversing this process.

The future development of the sizes of age group populations is in part determined by previous population history. It is relatively easy, for national populations, to show how the fertility fluctuations of the past are reflected in national age pyramids, and to examine the outputs of national projections to see the cohort-ageing process being worked out into the future.

Since 1911 (Rees and Warnes 1988, Rees 1992d) the UK population has been ageing in relative terms. The percentage of population aged 60 and over was only 8.1% in 1911 but had climbed to 20.2% by 1981. However, a naïve extrapolation of this trend over the past decade will not do: in 1991 only 20.8% were aged 60+ and in 2001 the percentage is projected to be 20.4. The ageing process had come to an end because the cohorts entering the elderly ages (here defined as 60 and over) were smaller in 1981-

2001 than in the previous twenty years (1961-81). The 1981-2001 cohorts entering old age were born in the inter-war years of low fertility.

However, before we assume that a happy equilibrium has been reached, it is necessary to break down the elderly into several component age groups and look at change for each of these. Figure 15 presents estimates of age group change for the UK population over the 1981-91 decade. A very varied pattern emerges. In particular, the age groups around retirement (60-64 and 65-74) show declines in all UK countries except Wales, but the gains in the middle elderly ages, 75-84, were substantial, around 16%, while those for the very elderly were spectacular, around 50%.

It is clear that we should not view old age as an inevitable period of dependency. Most old people are able to live fulfilling, independent lives but the probability of disability, limiting long-term illness and poverty rises steeply with age within the group. The family, community and health care needs of persons aged 80 and over are four to ten times those of persons aged 60 (Craig 1983). The National Health Service faces an increased demand on its services of perhaps 1% a year as a result of ageing within the elderly population.

What about the future evolution of the elderly age group? It is often assumed that, because fertility cannot affect

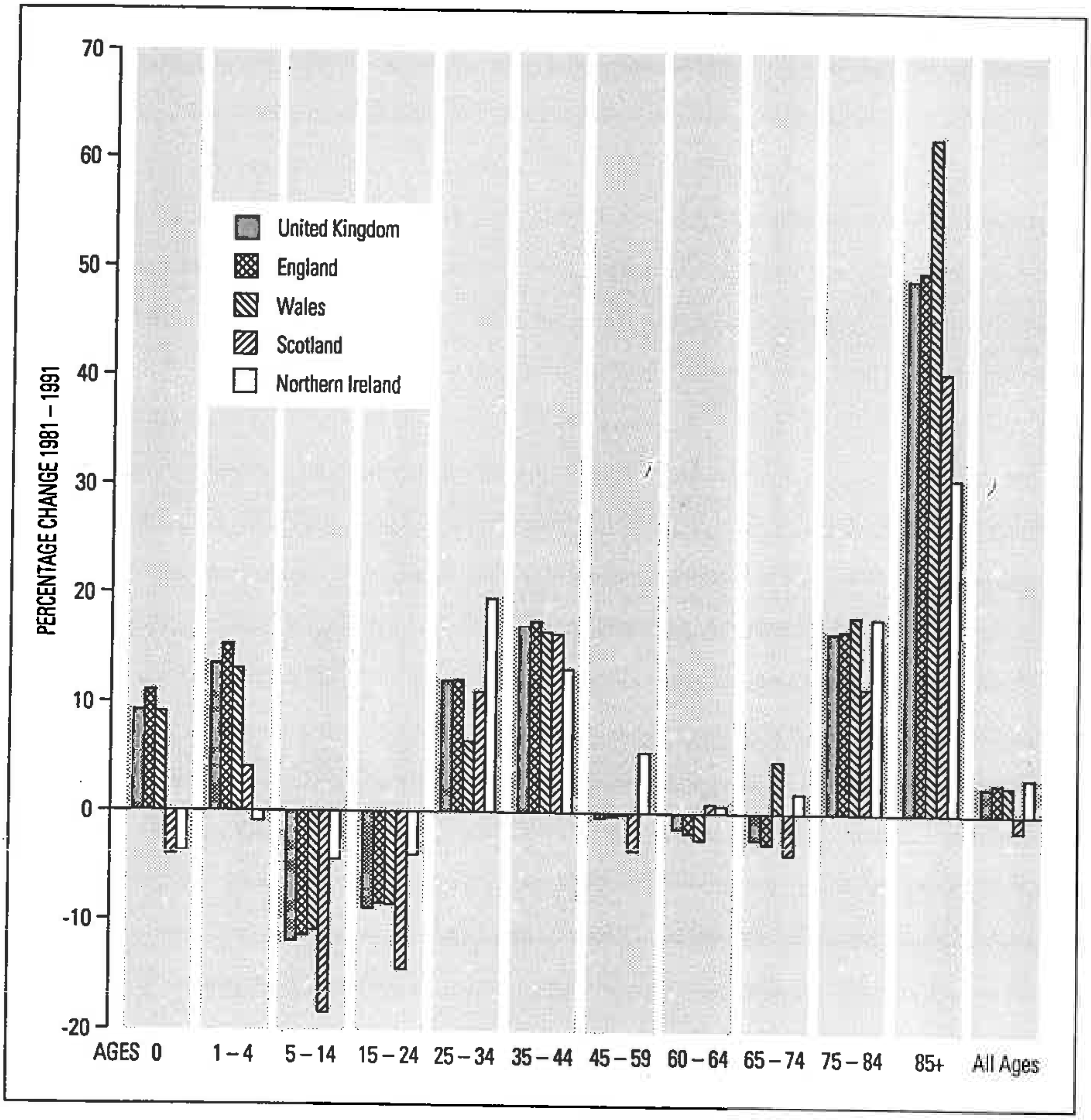

Figure 15: Age group change 1981 – 91, UK and countries

the projected size of the elderly population for 60 years and because elderly migration activity across international borders is relatively small in scale, we can be fairly certain about the future size of the national elderly population. However, it is clear that if we examine a suite of national projections based on different assumptions about the rate of improvement in mortality experience, that this is not the case.

Figure 16 compares two official projections from the 1980s (OPCS 1983 and 1989) and three variant projections (Rees 1992). OPCS (1983) adopts fairly conservative assumptions about mortality improvement (based on 1960s and 1970s experience) while OPCS (1989) builds in the improved performance of the 1980s (Figure 11). Variant A simply applies benchmark period (1981-86) fertility, mortality, internal and external migration rates without change until 2051. Variant B assumes that fertility rates rise to replacement level by 2000, that mortality rates decreased by constant percentages per quinquennium and that migration rates remain constant. Variant C adopts the same fertility and mortality assumptions as Variant B but assumes a higher rate of net immigration.

The top graph in Figure 16 shows that the growth of the elderly population will resume, under all scenarios, by 2001. Numbers of elderly will peak in 2031 as the baby boom cohorts become elderly and then decrease again

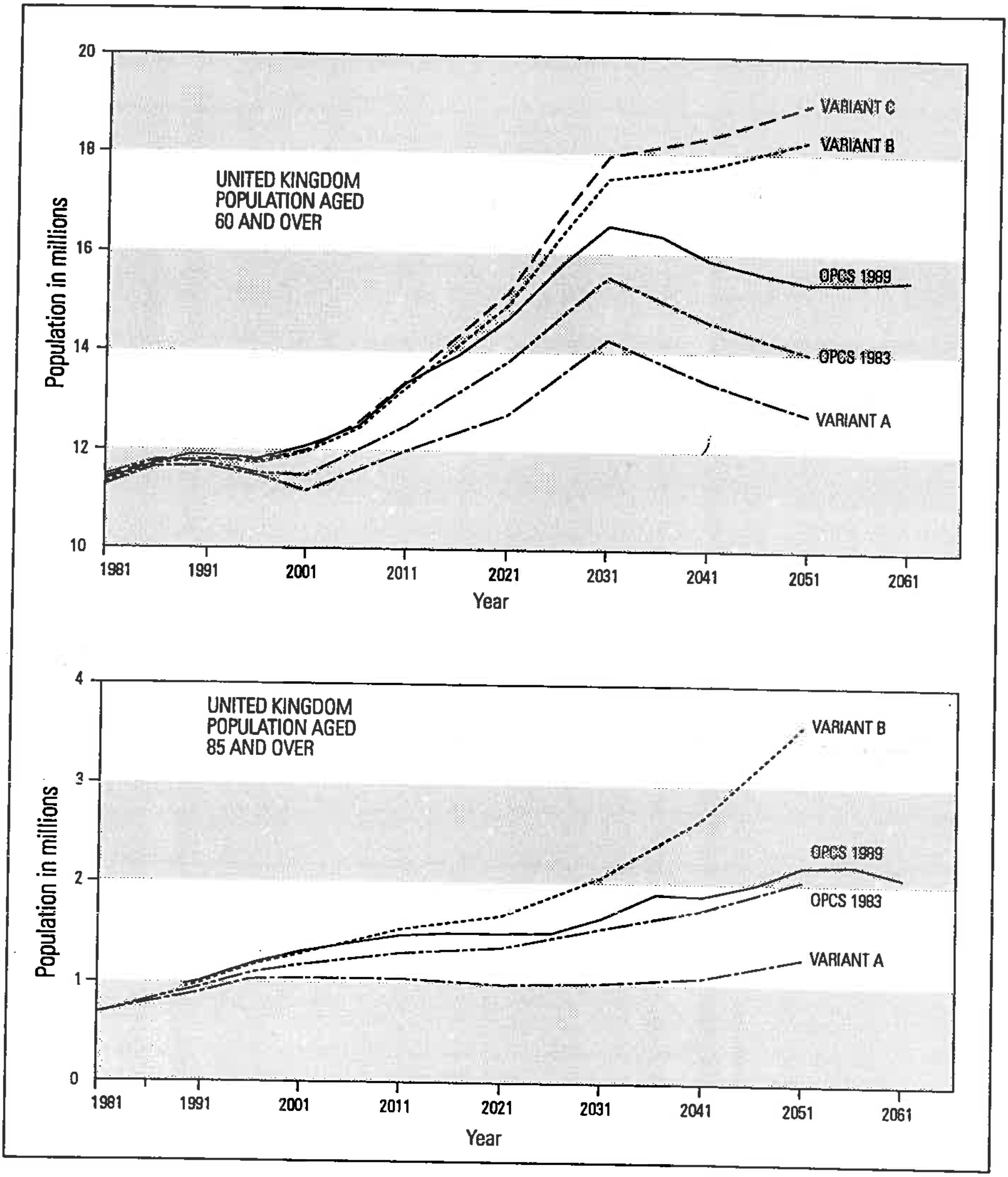

Figure 16: Variant projections of the numbers of elderly in the UK, 1981 – 2061

thereafter in the two official scenarios and under the constant assumption. However, both official projections assume a diminution and eventual cessation of mortality improvements. If instead we assume continuation of current improvement rates over the whole projection period, then the Variant B and C projections result in which increased survival changes outweigh the decreasing cohort effect after 2031. The life expectancies attained under Variants B and C (80 years for men and 85 for women) are equivalent to the most favoured contemporary populations, in the healthiest regions in Japan or Sweden.

One interesting implication of greater longevity will be the increasing frequency of the four generation family even though the generation age group is widening. Great grandparents will be in their 80s and 90s, grandparents in their 60s and 70s, parents in their 30s and 40s with children in pre-adult ages.

6. DEMOGRAPHIC ESTIMATION

The final topic which is treated in the paper is that of demographic estimation, the set of techniques which are used to bridge the gap between available data and the information required in the demographic model being employed. Full treatment of this topic is provided elsewhere (United Nations 1983, 1988; Rees and Woods 1986; Rees 1990, Chapter IV and V), so that here a few general principles and appropriate illustrations are outlined. Two common problems are addressed: how is the level of age disaggregation needed in the demographic problem to be achieved, and how is the level of spatial disaggregation desired to be attained?

The most useful first step is to undertake a data inventory for the estimation problem faced. Table 13 sets out a framework for such an inventory when the problem is achieving the detailed age classification required by a projection or life table model. The rows of the table set out the hierarchy of populations for which differing quantities of demographic data are available. The columns of the table arrange the types of variable from most detailed and relevant in the left most column to least detailed and relevant in the right most column. Both the model target variable represented by a (t) and the various kinds of available data (represented by a d) can be located in particular cells of the table. The task is then to make best use of the available data, under a set of simplifying

Table 13 A framework for conducting a data inventory for age classifications

Type of population/ spatial scale	Type of Variables: age disaggregation: 1 year ages V_1	5 year ages V_2	Broad ages V_3	All ages V_4	Indirect Variable V_5	Surrogate Variable V_6
Small subnational unit S_1						
Medium subnational unit S_2						
Large subnational unit S_3		t d(p)		d(m)		
National unit S_4						
Similar National unit S_5						
Model family population S_6		d(m)				

Notes

t = target variable needed for model

d = data available

(p) = for population

(m) = for migration

assumptions, to estimate the target variables. This might involve, for example, using a national set of mortality rates by quinquennial ages together with knowledge of county deaths by a set of broad ages to estimate the 5 year mortality rates needed for abridged life table construction. Such an estimation was employed in the course of producing the mortality statistics displayed on the earlier map of the UK (Figure 13).

Table 14 provides an illustration, for a developing country, of using data derived from a synthetic or model population to estimate age disaggregated out-migration rates for the country's capital region. The model population is the amalgamation of the migration experience of 17 developed countries, the migration age profiles of which were intensively studied by Rogers and Castro (1981). They fitted multi-exponential functions to the age profiles of over 200 inter-regional migration schedules. The model form and average model parameters are given in the Table. From these parameters are generated the standard model migration rates for the ages required in the Zimbabwe projection model. The model rates are multiplied by the average populations for Harare for the five year period prior to the 1982 Census to yield the number of migrants in each period cohort. The migrant numbers are summed and used to compute a crude migration rate. The model migration rates are then multiplied by the ratio of the model crude rate to an estimated crude rate derived from analysis of a lifetime migration table from the 1982 Census. The right most column in the table contains the result desired: the rates of out-migration of women from Harare to the Rest of Zimbabwe

Table 14 Migration rate estimates for females in Harare, Zimbabwe, 1977-82

Period-cohort Start 1977	End 1982	Average age	Standard model migration rate	Population at risk 1977-82 (1000s)	Estimated migration rate out of Harare
Birth	0- 4	0	0.0230	40.7	0.1073
0- 4	5- 9	5	0.0151	38.5	0.0706
5- 9	10-14	10	0.0104	34.4	0.0483
10-14	15-19	15	0.0075	39.4	0.0351
15-19	20-24	20	0.0278	39.3	0.1295
20-24	25-29	25	0.0364	32.5	0.1699
25-29	30-34	30	0.0257	21.9	0.1197
30-34	35-39	35	0.0170	15.0	0.0791
35-39	40-44	40	0.0115	11.3	0.0536
40-44	45-49	45	0.0081	8.7	0.0380
45-49	50-54	50	0.0061	6.6	0.0286
50-54	55-59	55	0.0048	5.2	0.0228
55-59	60-64	60	0.0041	3.7	0.0193
60-64	65-69	65	0.0037	3.5	0.0172
65-69	70-74	70	0.0034	2.8	0.0160
70-74	75-79	75	0.0033	2.0	0.0152
75-79	80-84	80	0.0032	1.1	0.0147
80-84	85-89	85	0.0031	0.4	0.0144
85+	90+	90	0.0031	0.3	0.0143
Totals			0.2174	307.1	CMR= 0.0847

Source: Rees (1991b)

Notes: 1. The standard model uses the average parameter values found by Rogers and Castro (1981, Table 17, p.42) in their study of over 200 inter-region profiles of migration rates by age. The model is:

m(x) =
a_1 exp (-alpha$_1$ x)
+ a_2 exp (-alpha$_2$ (x-mu$_2$) - exp (-lambda$_2$ (x-mu$_2$)))
+ C
where
a_1 = 0.02, alpha$_1$ = 0.10, a_2 = 0.06, mu$_2$ = 20, alpha$_2$ = 0.10, lambda$_2$ = 0.40 and c = 0.003

over a five year period. These rates are input to a two region projection model of Zimbabwe's population (Rees 1991b).

Additional spatial estimation problems are often faced when constructing multi regional projection models. Data may not be available for the full inter-regional origin-destination-age array of migration flows needed. Instead only partial information is easily acquired. Figure 17 shows the general situation involved. Data for some of the faces or edges of the array only may be available. Willekens, Por and Raquillet (1981) outline a variety of techniques for filling in the array. In the absence of critical faces or edges, we can instead employ one of the spatial interaction models of Wilson (1974) and so employ relationship between migration and distance to fill the data gaps.

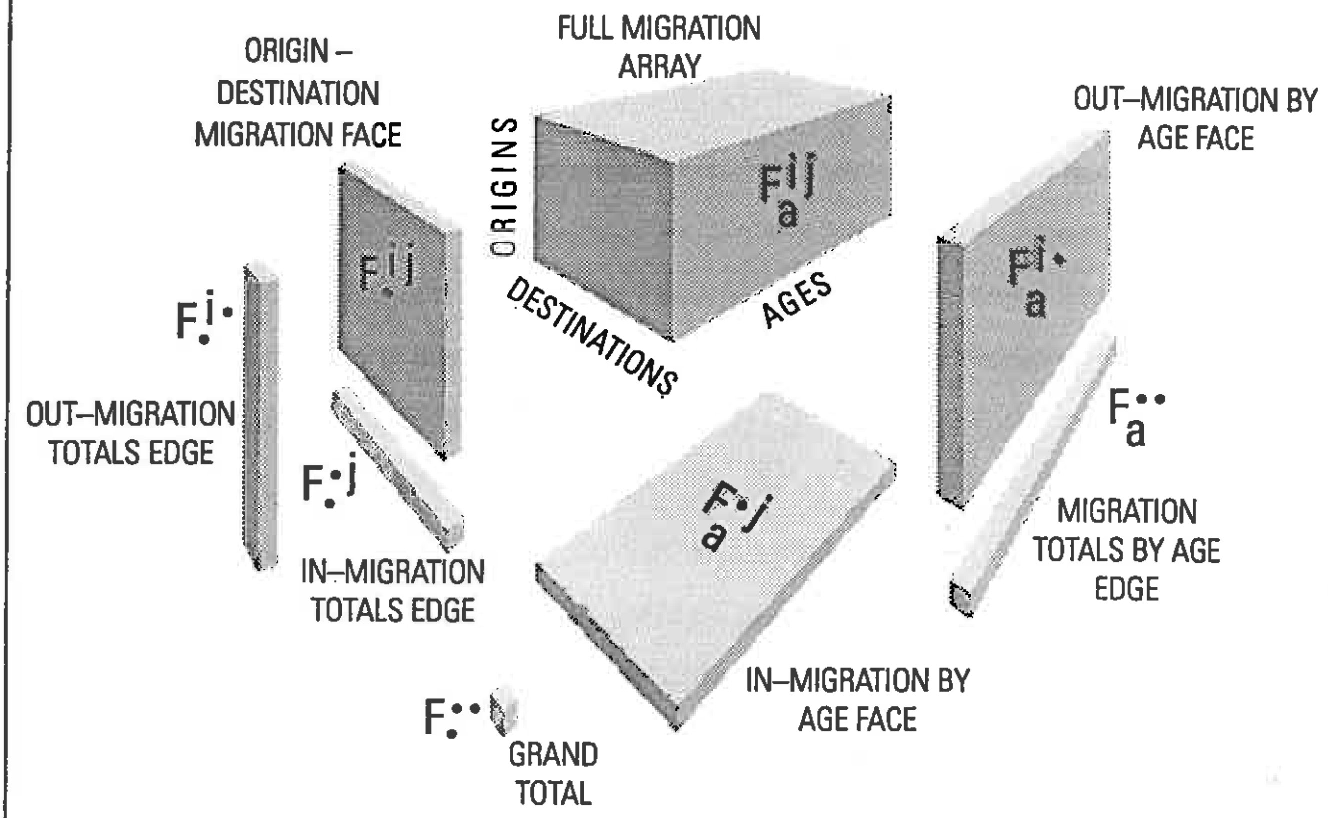

Figure 17: The migration cuboid, its faces and its edges

7. CONCLUSIONS: BEYOND POPULATION ARITHMETIC

The science of demography is, for the most part, concerned with the observation of the facts of population asking how many people are there, of what kinds and living in what places, asking how do these numbers change and why. However, population arithmetic alone is not enough.

In most systems of interest people have choices to make about consumption of goods, use of services, marriage, household formation, participation in the labour force, and many, many others. Knowledge of the numbers of people in socio-demographic detail is only one ingredient in understanding the behaviour of the system as a whole. In the middle of the 1980s both official and academic forecasts (DES 1984a, 1984b; Rees 1986b) of the number of students suggested that the number of students in Higher Education would not rise very much in the 1980s (from a base in the 1983-84 academic year) and would fall substantially in the 1990s as a consequence of the decline in the number of 17 to 19 year olds, reflecting the low numbers of births in the 1972 to 1982 period. These forecasts turned out to be very wrong for two reasons. The first reason was that the assumption that the age participation index - the proportion of 18 year olds qualified for and wishing to enter higher education - would remain constant proved untrue. The second reason was that the models underpinning the forecasts assumed that the numbers of students entering higher education were determined by demand, whereas in fact policy with regard to the supply and costing of places in Higher Education was crucial (Rees 1991c).

Population arithmetic alone is not enough for another reason. The population arithmetic is implicitly embedded in the prevailing value system of the researcher and of the society. The history of the ethnic question in the British Census is a case in point. In both 1971 and 1981 proposals for the introduction of such a question were opposed by both some politicians and some ethnic communities because the question was seen as threatening to individual liberty. During the 1980s, however, the view was accepted that, if there was to be effective monitoring of the socio-economic position of ethnic minorities in the 1991 Census, existing tabulations of the population by place of birth of the head of the household would be very inadequate. The question was asked with relatively little difficulty, after careful pre-testing.

There are many more examples of the interaction of our value systems and demographic investigations. But the most important point to make is to recognise the value of all the individuals that we count. Counting people in the past, present and future is vital because people themselves count.

REFERENCES

BECKER, K. (1874) *Zur Berechnung von Sterbetafeln an die Bevölkerungsstatistik zu Stellende Anforderungen*. Berlin, 76p. (Cited in Vanderschrick 1992).

BODEN, P., STILLWELL, J. and REES, P. (1991) Internal migration projection in England: the OPCS/DOE model examined. In Stillwell, J. and Congdon, P. (eds) *Migration models: macro and micro approaches*. Belhaven, London. Pp 262-286.

CEC, Commission of the European Communities, Directorate-General for Regional Policy (1991) *The Regions in the 1990s: Fourth Periodic Report on the Social and Economic Situation and Development of the Regions of the Community*. Office for Official Publications of the European Communities, Luxembourg.

CLARK, A. (1992) 1991 Census: collecting the data. *Population Trends*, 70, 22-27.

COMPTON, P. (1992) Migration trends for Northern Ireland: links with Great Britain. In Stillwell, J., Rees, P. and Boden, P. (eds) (1992) *Migration Processes and Patterns. Volume 2. Population Redistribution in the United Kingdom*. Belhaven, London. Pp 81-99.

COURGEAU, D. (1980) *Analyse Quantitative des Migrations Humaines*. Masson, Paris.

DALE, A. and Marsh, C. (eds) (1993) *The 1991 Census User's Guide*. HMSO, London.

DES (1984a) Demand for Higher Education in Great Britain, 1984-2000. *DES Report on Education Number 100*, July 1984.

DES (1984b) Technical Report to DES Report on Education 100. July 1988.

DEVIS, T. and MILLS, I. (1986) A comparison of migration data from the National Health Service Central Register and the 1981 Census. *Occasional Paper 35*, OPCS, London.

GREEN, A., WINNETT, C., HASLUCK, C. and OWEN, D. (1993) Recent unemployment change: spatial perspectives. Institute for Employment Research, University of Warwick. Paper presented at Annual Conference of the Institute of British Geographers, Royal Holloway and Bedford New College, University of London, Egham, Surrey, January 1993.

HAVERKATE, R. and VAN HASELEN, H. (1990) *Demographic Evolution through Time in European Regions (DEMETER 2015)*. Report to the Commission of the European communities, Directorate-General for Regional Policy, Netherlands Economic Institute, Rotterdam.

HMSO (1992) *The 1992/93 Publications Catalogue. Office of Population Censuses and Surveys*. HMSO, London.

HOWE, M. (1963) *National Atlas of Disease Mortality in the United Kingdom*. Nelson, London.

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

LEXIS, W. (1975) *Einleitung in die Theorie der Bevölkerungsstatitistik*. Trübner, Strasbourg.

OPCS (1992a) 1991 Census, Great Britain National Monitor. Census 1991 CM56. OPCS, London.

OPCS (1992b) *1991 Census. Definitions*. Cen 91 DEF. HMSO, London.

OPCS (1992c) *Populations Trends 70*. HMSO. London.

PRESSAT, R. (1961) *L'Analyse Démographique*. Presses Universitaires de France, Paris. 402p.

REES, G. and REES, P. (1990) *Programmers' documentation. SWIS: Swansea Information System. A ward population information system for Swansea City Council*. GMAP Ltd, University of Leeds.

REES, P. (1977) *The future population of East Anglia and its constituent counties (Cambridge, Norfolk and Suffolk)*. A report prepared for the East Anglian Economic Planning Council.

REES, P. (1981) Accounts based models for multiregional population analysis: methods, program and user's manual. *Working Paper 295*, School of Geography, University of Leeds.

REES, P. (1984) Spatial population analysis using movement data and accounting methods: theory, models, the 'MOVE' program and examples. *Working Paper 404*, School of Geography, University of Leeds/*Working Paper No.50*, Netherlands Inter-Disciplinary Institute, The Hague, Netherlands.

REES, P. (1985) Does it really matter which migration data you use in a population model? Chapter 5 in White, P. and Van der Knapp, B. (eds) *Contemporary Studies in Migration*. Geo Books, Norwich. Pp 55-77.

REES, P. (1986a) Choices in the construction of regional population projections. In Woods, R. and Rees, P. (eds) *Population Structures and Models*. Allen and Unwin, London. Pp 126-159.

REES, P. (1986b) A geographical forecast of the demand for student places. *Transactions, Institute of British Geographers*, NS11, 5-26.

REES, P. (1989a) *Britain's Population: a Geographical Analysis*. Draft Manuscript, School of Geography, University of Leeds.

REES, P. (1989b) 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. (1990) *Manual for Subnational Population Projections*. Draft manuscript prepared for the Population Division, Department of International Economic and Social Affairs, United Nations, New York. School of Geography, University of Leeds.

REES, P. (1991a) The projection of small area populations: a case study in Swansea, Wales. Paper presented at the International Seminar on Population Geography, Soesterberg, the Netherlands, 7-10 April 1991. Forthcoming in *Dutch Geographical Studies*, Utrecht and Amsterdam.

REES, P. (1991b) A tale of two cities: the prospects for future ageing of the populations of London and Harare. In Department of International Economic and Social Affairs, *Ageing and Urbanisation*. United Nations, New York. Pp 117-166.

REES, P. (1991c) The future of the student in higher education: numbers, flows, competition and planning in a geographic context. *Working Paper 91/2*, School of Geography, University of Leeds.

REES, P. (1991d) A projection model for small area populations: applications to West Yorkshire Districts. *Working Paper 91/10*, School of Geography, University of Leeds.

REES, P. (1992a) The ESRC/UFC-ISC 1991 Census of Population initiative: delivering the data of the decade. *ESRC Data Archive Bulletin*, 51, 12-22.

REES, P. (1992b) Projecting the populations of urban communities. Paper presented at the US/UK Seminar on Migration in Post Industrial Societies, University of Southern California. August 1992. Submitted to *Environment and Planning A*.

REES, P. (1992c) Elderly migration and population redistribution in the United Kingdom. In Rogers, A. (eds) *Elderly Migration and Population Redistribution*: a Comparative Perspective. Belhaven, London. Pp 203-225.

REES, P., STILLWELL, J. AND CONVEY, A. (1992) Intra-Community migration and its impact on the demographic structure at the regional level. *Working Paper 91/1*, School of Geography, University of Leeds.

REES, P. and WARNES, A. (1988) United Kingdom. In Rogers, A. and Serow, W. (eds) *Elderly Migration: an International Comparative Study*. Population Program, Institute of Behavioural science, University of Colorado, Boulder, Colorado.

REES, P. and WOODS, R. (1986) Demographic estimation: problems, methods and examples. In Woods, R. and Rees, P. (eds) *Population Structures and Models*. Allen and Unwin, London. Pp 301-343.

REES, P. and WILSON, A (1977) *Spatial Population Analysis*. Edward Arnold, London.

ROGERS, A. (1976) Shrinking large scale population projection models. *Environment and Planning A*, 8, 515-541.

ROGERS, A. (1986) Parameterized multistate population dynamics and projections. *Journal of the American Statistical Association*, 81, 48-61.

ROGERS, A. and CASTRO, L. (1981a) *Model Migration Schedules*. Research Report RR-81-30. International Institute for Applied Systems Analysis, Laxenburg, Austria.

ROGERS, A. and CASTRO, L. (1981b) Age patterns of migration: cause-specific profiles. In Rogers, A. (ed) *Advances in Multiregional Mathematical Demography*. Research Report RR-81-6. International Institute for Applied Systems Analysis, Laxenburg, Austria. Pp 125-159.

ROGERS, A. and PLANCK, F. (1983) MODEL: a general program for estimating parameterized schedules of fertility, mortality, migration, marital and labour force status transitions. *Working Paper WP-83-1-2*, International Institute for Applied Systems Analysis, Laxemburg, Austria.

SASPAC (1992) *SASPAC User Manual, Parts 1 and 2*. CST 450. Produced at Manchester Computing Centre in accordance with an agreement between the Combined Higher Education Software Team (CHEST) and the London Research Centre.

SHEN, J. (1991) Analysis of urban-rural population dynamics for China. *Environment and Planning A*, 23, 1797-1810.

SHEN, J. and SPENCE, N. (1993) Population trends, labour supply and the employment problem in China. Department of Geography, London School of Economics and Political Science. Paper presented at the Annual Conference of the Institute of British Geographers, Royal Holloway and Bedford New college, University of London, Egham, Surrey, January 1993.

STILLWELL, J., REES, P. and BODEN, P. (eds) (1992) *Migration Processes and Patterns. Volume 2. Population Redistribution in the United Kingdom.* Belhaven, London.

STONE, R. (1971) *Demographic Accounting and Model-building.* Organisation for Economic Co-operation and Development, Paris.

TICKELL, A. (1993) End of the great divide as South-east is dragged into the jobless mire. *The Guardian*, 4 January 1993, p 11.

UNITED NATIONS (1970) *Manual VI: Methods of Measuring Internal Migration.* ST/SOA.Series A/47. United Nations, New York.

UNITED NATIONS (1976) *Towards a System of Social and Demographic Statistics.* United Nations, New York.

UNITED NATIONS (1983) *Manual X: Indirect Techniques for Demographic Estimation.* Department of International Economic and Social Affairs. Population Studies, No. 81. United Nations, New York.

UNITED NATIONS (1988) *Mortpak - the United Nations Software Package for Mortality Measurement.* Department of International Economic and Social Affairs. United Nations, New York.

VANDERSCHRICK, C. (1992) Le diagramme de Lexis revisité. *Population*, 5, 1241-1262.

VERWEIJ, A. (1974) *De Waarnemingen der Bevolkingsstatistiek.* Academisch Proefschnift, ter verkrijging van den graad van doctor in de wis - en natuurkunde, aan de Hoogschool te Utrecht. (Cited in Vanderschrick 1992).

VERWEY, A. (1975) Principles of vital statistics. *Journal of the Statistical Society*, XXXVIII, 487-513. (Cited in Vanderschrick 1992).

WILLEKENS, F. and DREWE, P. (1984) A multiregional model for regional demographic projection. In ter Heide, H. and Willekens, F. (eds) *Demographic Research and Spatial Policy.* Academic Press, London. Pp 309-334.

WILLEKENS, F., POR, A. and RAQUILLET, R. (1981) Entropy, multiproportional and quadratic techniques for inferring patterns of migration from aggregate data. In Rogers, A. (ed) *Advances in Multiregional Mathematical Demography.* Research Report RR-81-6. International Institute for Applied Systems Analysis, Laxenburg, Austria. Pp 83-124.

WILLEKENS, F. and ROGERS, A. (1978) *Spatial Population Analysis: Methods and Computer Programs.* Research report RR-78-18. International Institute for Applied Systems Analysis, Laxenburg, Austria

WILLIAMS, J. (1992) A simulation model for the spread of HIV and AIDS in the United Kingdom. *Working Paper 92/2*, School of Geography, University of Leeds.

WILLIAMS, J. (1993) Adding drug users to a model of sexual and vertical transmission of HIV in the UK. Research Support Group Report Numbers, School of Geography, University of Leeds, January 1993.

WILSON, A. (1974) *Urban and Regional Models for Geography and Planning.* Wiley, Chichester, Sussex.

WILSON, C. (ed) (1985) *Roland Pressat's Dictionary of Demography.* Blackwell, London.

ABBREVIATIONS

DES = Department of Education and Sciences
HMSO = Her Majesty's Stationery Office
OPCS = Office of Population Censuses and Surveys