WORKING PAPER 486

A MODEL FOR UPDATING INDIVIDUAL AND HOUSEHOLD POPULATIONS

Philip REES

Martin CLARKE

and

Chris DULEY

School of Geography
University of Leeds
Leeds LS2 9JT

March 1987

CONTENTS

ABSTRACT

This paper provides a first sketch of a model designed to effect the updating of small area populations in the time interval between British censuses. Several further sketches must follow this one, however, before the final picture of Britain's population in 1987 or 1988 can be painted.

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

This research is supported by a Collaborative Award in the Social Sciences made available by the Economic and Social Research Council to support a postgraduate studentship at the School of Geography in collaboration with Pin Point Ltd of London, who provide further financial assistance. Our thanks to both ESRC and Pin Point.

of simulated individuals and households with their interesting characteristics, and then evolving those individuals over time to the present using "fresh" information about how those attributes change over time. The updating process would be taken forward in time to the next census both to provide a set of forecasts and to make possible a test of methods in 1991/92. Additionally non-census variables, such as income, taxation, benefits, expenditure and so on, may also be incorporated into the analysis.

The general structure of the model is set out in Figure 1. The model is divided into three phases. The first, Population Reconstruction, involves simulating a list of individual, household and housing attributes at a base point in time, using as much aggregate information as can be incorporated. The second phase involves updating the numbers in each population and the characteristics of each unit, employing survey and list data for years following the base date. Data for updating are available from records of demographic events at national, regional and local scales, from household surveys at national and regional scales and from administrative records such as address lists of various kinds at local scales. The third phase of the model will involve projecting changes in population numbers and characteristics into the future using trended event data and attribute transition data together with a range of alternative scenarios or policies (conditional forecasting). The projection phase would last until the next census, when thorough evaluation of the whole model could be carried out, and the lessons incorporated into a revised model based on the next census.

3. POPULATION RECONSTRUCTION

To reconstruct the population decisions are needed about attributes, about how these can be estimated from data in the Small Area Statistics and elsewhere, and about the sequence of probability equations that are to be used.

3.1 The attribute list

These can be divided into four sets: those for individuals, for households, for housing units and for context attributes. These are listed in turn. Each attribute is given an index number, a short title, a number of categories and the labels for those categories. Some attributes may be continuous in form rather than categorical.

3.1.1 Individual attributes

(I1)	status within household (2)	- head, non-head
(I2)	status within family (5)	- family head, spouse, dependent child, non-dependent child, other person (non-family member)
(I3)	age (C) (77)	- single years to 75, 75 and over
(I4)	sex (2)	- male and female
(I5)	marital status (4)	- single, married, divorced, widowed
(I6)	ethnic status (13)	- origin: British, Irish, AuCaNz, Pakistani, Indian, Bangladeshi,

1. INTRODUCTION

This paper sets out ideas about how a model for updating populations of individuals and households should be constructed. The model will start with information from the latest census (Census 1981), reconstruct socioeconomic profiles of the population and then update this information for years up to the present and then on to the next census year, 1991, using microsimulation methods. The justification for adopting a microsimulation approach for population updating in preference to other methods set out in Clarke (1986). Briefly, the main advantages are:

(1) Efficiency of a micro-level representation of complex systems characterised by interdependence and heterogeneity in terms of computer storage.

(ii) The appeal of modelling demographic processes at the level at which they take place. This is particularly important in the case of processes, such as marriage and divorce which have important consequences for household formation, disolution and fission.

(iii) The use of continuous variables, such as age and income without the necessity to aggregate into groups.

(iv) Flexible aggregation of the results to produce selected tabulations.

For examples of previous work on the use of microsimulation methods in population studies the reader is referred to the reviews in Clarke (1986) and Clarke and Holm (1984).

The paper builds on discussions over the past five months at the School of Geography and progress made by Chris Duley (Duley, 1987). The general shape of the model must be specified before detailed implementation can be undertaken. However, the detailed structure of the model is likely to continue to evolve substantially as more is learnt about possible sources of updated data and about the feasibility of particular features as the model is converted into a computer program (as discussed in Clarke and Holm, 1987).

2. THE AIMS OF THE MODEL

Currently, both public and private sector decision makers continue to use 1981 Census data in situations where geographical detail is essential - targetting of area based policies to alleviate deprivation or targetting of direct mail marketing campaigns. However, it rapidly dates both nationally as a result of social and economic trends, and through local processes of housing demolition, decay, redevelopment and construction, and the associated population processes of mobility and turnover.

The broad aim of this proposed model is to improve on this geographic database by updating it. This will be achieved by taking a base population of individuals and households measured at a previous census (1981) in a small area, reconstructing it as a list

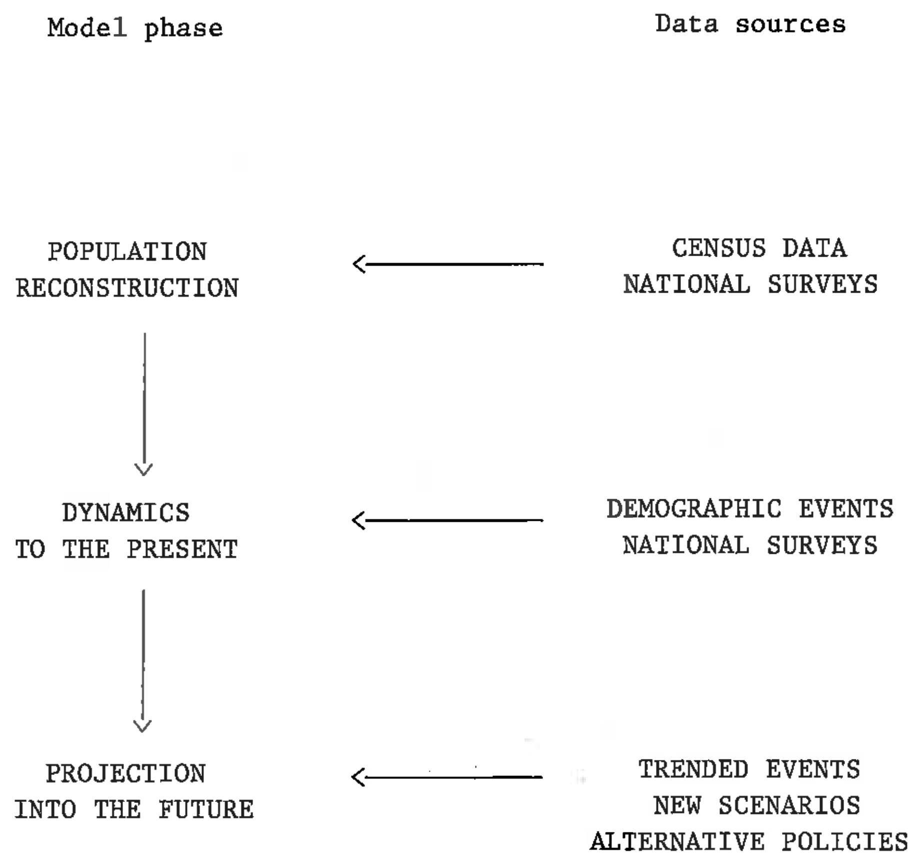


FIGURE 1. <u>The general structure of the updating model</u>

		West Indian, Asian from E.Africa, African, NC Remainder, Other E.C., Other European, Rest of the World
(I7)	place of birth (2)	- inside UK, outside UK
(I8)	education achieved (5) (highest qualification)	- none, CSEs, O levels, A levels, degree or professional
(I9)	economic activity (4)	- active, retired, student, inactive
(I10)	socioeconomic group (18)	- s.e.g.s I to XVII with V split into V.1 and V.2
(I11)	employment status (4)	- full-time, part-time, self-employed, unemployed
(I12)	industry (7)	- agriculture forestry and fishing energy and water manufacturing construction distribution and catering transport and communication other services
(I13)	earned income (C)	- income by virtue of work

In addition individuals would need to be allocated reference numbers to facilitate the list processing:

(IR1)	household number	- sequence number of household generated
(IR2)	family number	- family number within household (if no family were present, this would be set to zero)
(IR3)	individual number	- individual number within family

There would be a separate list for individuals living in non-private households. The same characteristics would be reconstructed with the following exceptions:

(IN1)	establishment status	- staff, residents
(IN2)	establishment type (9)	- hotel/boarding houses, children's homes, old people's homes, psychiatric hospitals, other hospitals, schools & colleges, prison dept. estabs., lodging houses, other estabs.

The reference numbers would be

(INR1)	establishment number	- sequence number of establishment (at present no such information has been traced and so this would be set to 1
(INR2)	individual number within establishment	- persons number within establishment

3.1.2 Household attributes

The following list is suggested for households:

(H1) type based on structure - no family households:

and composition (10)

(A) one person
(B) two or more persons
<u>one family households, lone parent</u>:
(C) non-dependent child(ren) only, with other(s)
(D) non-dependent child(ren) only, without other(s)
(E) non-dependent and dependent child(ren), with other(s)
(P) non-dependent and dependent child(ren), without other(s)
(G) dependent child(ren) only, with other(s)
(H) dependent child(ren) only, without other(s)
<u>one family households, married couples</u>:
(I) no child(ren), with other(s)
(J) no child(ren), without other(s)
(K) non-dependent children only, with other(s)
(L) non-dependent child(ren) only, without other(s)
(M) non-dependent and dependent child(ren), with others(s)
(N) non-dependent and dependent child(ren), without others(s)
(J) dependent child(ren) only, with other(s)
(P) dependent child(ren) only, without other(s)
<u>two or more family households</u>:
(Q) two or more families

This would not be generated as an independent attribute but would be a product of the reconstruction of the initial population. The proposed scheme is outlined later (section 3.2). Tables of household structure and household size would be used as checks or constraints in the estimation process. The definition of dependency is in part based on age and in part on whether, for children aged 16 or over, they are in full time education.

(H2) number of persons in the household (7) - 1, 2, 3, 4, 5, 6, 7 or more This would be derived in the population reconstruction process.

(H3) earned income (C) - continuous variable which is the sum of incomes earned by the individual employed members of the household

(H4) benefits (C) - continuous variable which is the sum of benefits received by individual members (a product of applying benefit rules)

(H5) unearned income (C) - income from savings and investments

(H6) taxation (C) - taxes, the result of applying tax rules to H3, H4 and H5

(H7) net household income (C)	- the result of H3+H4+H5-H6
(H8) tenure (6)	- owner occupied rented from council or new town housing association rented by virtue of business or job privately rented, furnished privately rented, unfurnished
(H9) rooms (7)	- 1, 2, 3, 4, 5, 6, 7+
(H10) amenities (4)	- exclusive use of all amenities lacking indoor w.c. lacking hot water lacking indoor w.c. and hot water
(H11) value or rent (C)	- either the value of the housing unit or rental paid for the unit
(H12) number of cars (3)	- no car, one car, two or more cars

In addition household reference numbers will be needed:

(HR1) household number	- household sequence number

3.1.3 <u>Housing unit variables</u>

(U1) occupance (3)	- occupied first home, occupied second home, vacant
(U2) tenure (6)	- owner occupied rented from council or new town housing association rented by virtue of business or job privately rented, furnished privately rented, unfurnished
(U3) number of rooms (7)	- 1, 2, 3, 4, 5, 6, 7+ This would be generated using probablities dependent on size of household and tenure
(U4) amenities (4)	- exclusive use of all amenities lacking indoor w.c. lacking hot water lacking indoor w.c. and hot water
(U5) value or rent (C)	- either the value of the housing unit or rental paid for the unit

3.1.4 <u>Context variables</u>

These would not be generated directly in the microsimulation but could be derived after each year's population had been reconstructed.

(C1) PIN classification	- PIN POINT neighbourhood class
(C2) PIN wealth indicator	- PIN POINT derived index

These are constructed from the whole reconstructed population.

Examples of what the generated lists would look like are provided in Table 1 for individuals, Table 2 for households and Table 3 for housing

TABLE 1. An example of a database of individual attributes

Sequence numbers	Reference numbers		Status in Household, family	Age	Sex	Marital status	Ethnic status	Place of birth	Education level	Economic activity	Emp. status	S.E.G.	Industry	Earned income
	Household	Individual	I1, I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	I13
1	1	1	head	32	M	M	1	U	D	ACT	F	1	7	12K
2	1	2	spouse	29	F	M	1	U	A	ACT	P	2	7	7K
3	1	3	child	7	M	S	1	U	N	STU	-	-	-	-
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1999	713	5	child	2	F	S	3	A	N	OIN	-	-	-	-
2000	714	1	head	60	F	W	1	U	N	RET	-	-	-	-

TABLE 2. An example of a database of household attributes

Sequence number	Reference numbers		Type	Size	Earned income	Bene-fits	Un-earned income	Ta-xes	Net income	Number of cars	Tenure	Rooms	Amenities
	House-hold	Family	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11
1	1	1	7	3	19K	.4K	1K	5K	14.4K	1	0	7	ALL
2	2	1	2	2	25K	0K	0K	6K	19K	2	0	6	ALL
3	3	1	1	1	7K	0K	0K	1K	6K	0	F	3	WC
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
713	713	2	8	8	12K	3K	0K	2K	11K	0	0	5	ALL
714	714	1	1	1	.5K	2K	0K	0K	2.5K	0	0	2	ALL

TABLE 3. An example of a database of housing unit attributes

Sequence	Reference numbers Household	Occupance U1	Tenure U2	Rooms U3	Amenities U4	Value or rent U5
1	1	0	0	7	ALL	50K
2	2	0	0	6	ALL	25K
3	3	0	F	3	WC	10K
⋮	⋮	⋮	⋮	⋮	⋮	⋮
749	-	V	0	7	ALL	40K
750	-	V	C	6	ALL	28K

units.

3.2 Steps in the population reconstruction

3.2.1 Ordering of the reconstruction

The attributes in the lists above have been arranged in the order in which they will be generated. Individual characteristics will be generated within a family and a household loop in the way indicated in Figure 3. In these loops the first two attributes are created. Then, the attributes of each individual will be generated in turn from I3 through I13, in part dependent on the attributes of other members of the family and household whose generation has already occurred. The individual attributes are arranged in rough socioeconomic life course order. Characteristics acquired at birth - sex, ethnic status and place of birth - are generated early in the order and do not change subsequently, along with age which changes in an immutable fashion. Marital status is also defined early in the sequence because of its close association with the prior attributes, although, of course, it will tend to change systematically through the life course. Then follows the first attribute achieved in the life course, that of educational qualifications. This is in turn followed by four variables linked to activity in the economy which together generate earned income.

The household characteristics are firstly summations of attributes of the individuals that make up the household (H1 to H7) and then four consumption variables. The first three of these (H8 to H11) concern the consumption of shelter and the final household attribute H12 details one of the most important items of household expenditure after housing, the car. Many other consumption characteristics could be added to the household list such as a range of consumer durables or non-durables. This will not be done in the initial model prototype, but care will be taken to make it easy to plug in such characteristics at a later stage.

The housing unit attributes (U2 to U5) parallel those for housing consumption by households (H8 to H11) but the first attribute (U1) indicates whether the unit is occupied or not, and when occupied, whether as a first or second home. The characteristics of vacant housing units will need to be reconstructed separately from those of occupied units.

3.2.2 Generating heads and other members

These steps define the number of individuals whose socioeconomic profiles have to be simulated. The head's characteristics would be simulated first, and then those of the other members of the household. But a robust algorithm ("a series of instructions or procedural steps for the solution of a specific problem" - Chandor with Graham and Williamson, 1985, p.31) must be designed.

The number of heads in the reconstructed population will simply be the number of heads given in the SAS tables for the locational unit to be used (the postcode sector - see section 4). The number of heads is equal to the number of households in which at least one of the usual residents was present on census night. Adding in vacant and visitor households would tend to result in double counting.

How are non-heads to be simulated? The method will depend rather critically on the household classification adopted. A good basis on which to start work is the classification used in the 1981 Census National tables and in Census Guide 4 on Households (OPCS and COI, 1986, Figure 3, p.4); the ten categories in this classification are listed under household attribute H1 above. The classification is represented diagrammatically in Figure 2.

Private households are first classified by the number of families they contain, either none, one or two or more. A family consists of (OPCS, 1984, p.ix-x)

"a) a married couple with or without their never-married child(ren) (includes childless married couples),
or
b) a father or mother together with his or her never-married child(ren),
or
c) grandparent(s) with grandchild(ren) if there are apparent parents of the grandchild(ren) usually resident in the household."

No family households contain either persons living alone (just over one fifth of all private households) or unrelated persons living together (where relatedness refers only to the family as defined above). One family households are divided firstly into those with lone parents and those with married couples, and secondly according to the presence or absence of non-dependent children, dependent children or both, where, according to OPCS (1984, p.x)

"Dependent children are children in families who are

a) under 16 years of age, or
b) under 19 years of age, never-married and classified from the question on economic activity last week as a student."

Finally, households with more than one family are lumped together. They constitute only 1% of all households though a higher percentage of persons will live in this kind of household. Note that any household may contain additional non-related members, although many of these members will be either relatives in the wider sense or partners in consensual unions.

It follows from this classification of private households that the reconstruction of the population should proceed as a series of loops, as set out in Figure 3. The first loop generates the rquired number of households and within the loop the required number of families and non-family individuals. The second loop generates individuals living in non-private households (institutions), and the third loop generates vacant housing units.

				Numbers of households in millions	% of total	Type
All households/heads	No family		1 person	4.2	21.4	H1-A
			2+ persons	1.0	5.1	H1-B
	One family	Lone parent	Non-dependent children only	0.7	3.6	H1-C+H1-D
			Dependent and non-dependent children	0.2	1.0	H1-E+H1F
			Dependent only	0.7	3.6	H1-G+H1H
		Married couples	No children	5.0	25.5	H1-I+H1J
			Non-dependent children only	1.6	8.2	H1-K+H1L
			Dependent children and non dependent	1.2	6.1	H1-M+H1N
			Dependent children only	4.8	24.5	H1-O+H1P
	Two or more families			0.2	1.0	H1-Q
			Total, Great Britain	19.5	100.0	

Source: OPCS and COI (1986), p. 4.

FIGURE 2. <u>The household classification</u>

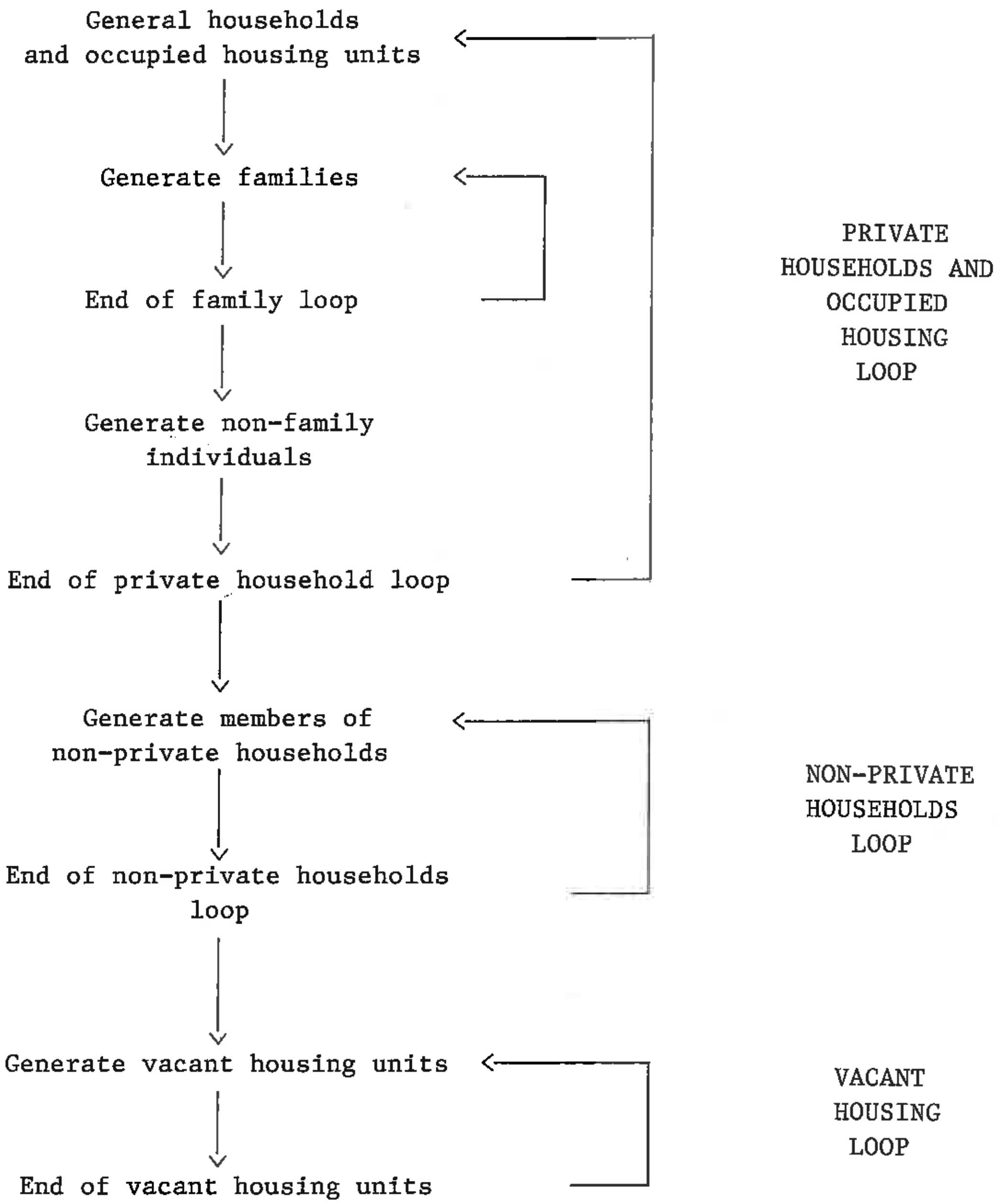


FIGURE 3. The general loop structure of the population and housing reconstruction process

The detailed structure of the private household loop is shown in Figure 4. Within the family loop is a spouse loop which is notional only under British law (but which might be relevant in polygamous or polyandrous societies). Within the family and spouse loops are separate dependent child and non-dependent child loops. Outside the spouse loop is a loop for generating other members of the family household. Outside the family loop is a loop for generating members of non-family households.

3.2.3 Constraining the reconstruction to SAS constraints

Since the model is designed to reconstruct the whole population of a small area, the Small Area Statistics (SAS) in the 1981 Census can be used as constraints. This means that the number of households, families or individuals with particular attributes should add up to the number observed in the relevant SAS table.

One approach to applying such constraints is depicted in Figure 5. Once a household and its attributes and all its members and their attributes have been generated, the relevant counts of households and individuals are increased by one. These counts are then compared with the known SAS counts. If the SAS count has been exceeded then the category is closed, and the reconstructed household is discarded and the household loop continues. This is merely an extension of the "end of loop index", which will be the number of households observed in the 1981 SAS.

A possible problem with the above procedure is that towards the end of the process it may take many passes through the household loops before an acceptable household is found. If that poses computational difficulties, then the process may need modification. When household categories are "closed", the probability distributions being used in the reconstruction are adjusted. An example of how this might operate is worked through in Table 4.

Assume we have figures for heads by age and sex, and thus a probability distribution by age and sex which is sampled via a random number function. We have to create 500 heads and generate one random number for each head, which is used to assign a head to an age-sex catogory. The number 82 would cause us to assign the head to the "female, 35-44" category. We keep count of the number of assignments to each category, and when one becomes full, we "close" it and recompute the probablities. This process continues until all heads have been created and all categories closed. This technique can be applied to all attributes for which we have SAS data. In one sense, this is a "learning" procedure - the model adjusts its "behaviour" on the basis of past actions.

3.3 Ordering of attributes and linking of the probabilities

The order in which the attributes will be simulated has been provisionally set out in section 3.1. A final decision will depend on a precise description of all the crosstabulations available in the SAS, in county or national volumes, together with tables from selected national household surveys. This description is currently being assembled. From these

Start of private household loop

Generate heads of household and attributes

Start of family loop

Generate head of family and attributes

Start of spouse loop

Start of dependent child loop

Generate dependent child and attributes

End of dependent child loop

Start of non-dependent child loop

Generate non-dependent child and attributes

End of non-dependent child loop

End of spouse loop

Start of other members of family household loop

Generate other members and attributes

End of other members of family household loop

End of family loop

Start of members of non-family household loop

Generate members and attributes

End of members of non-family household loop

End of private household loop

FIGURE 4. The loop structure for private households

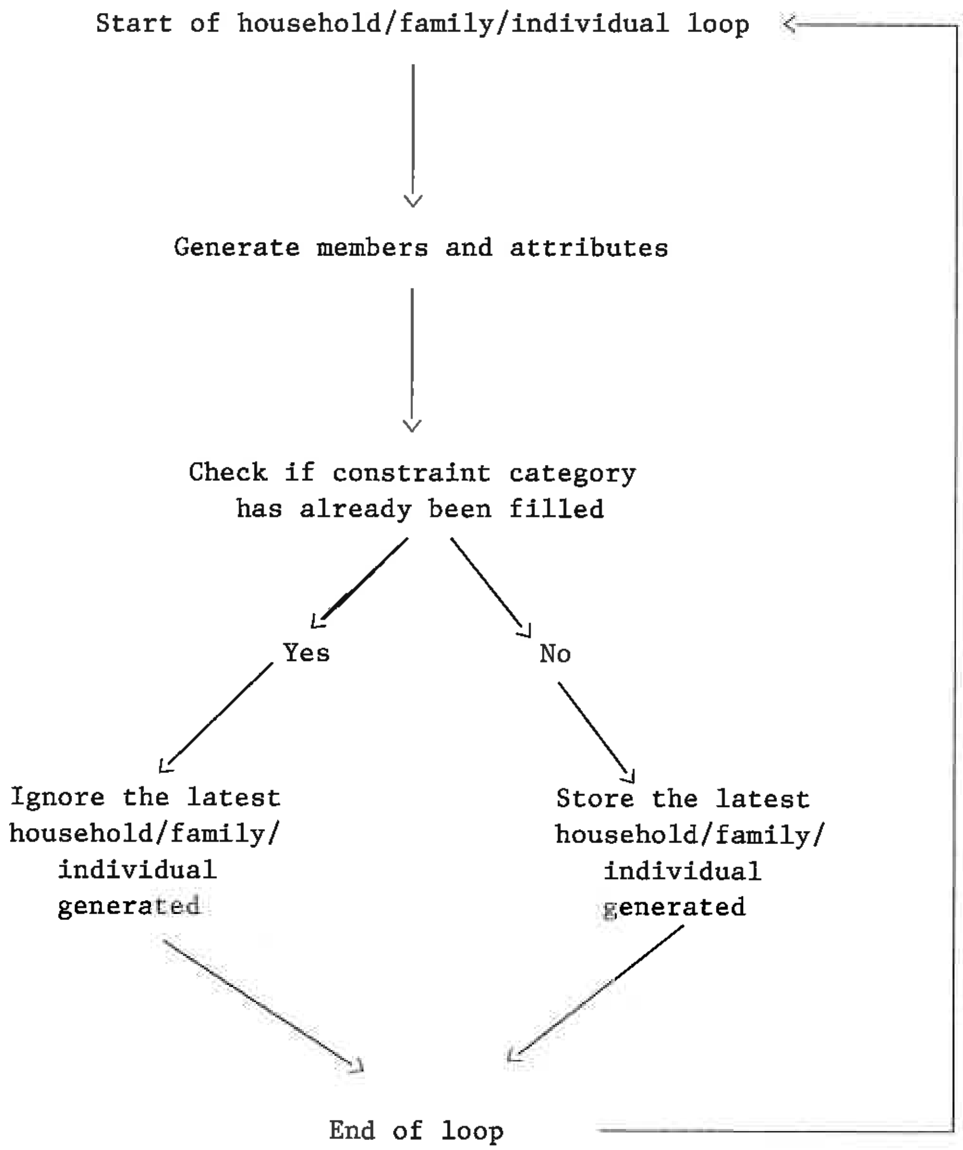


FIGURE 5. Constraint checking in the population reconstruction process

TABLE 4. Constraint calculations

Category	Sex-age category		Number	%	Cumulative %	When category 7 is closed	Cumulative %
1	Male	16-24	20	4	4	4.17	4.17
2		25-34	20	4	8	4.17	8.34
3		35-44	200	40	48	41.67	50.00
4		45 +	100	20	68	20.83	70.83
5	Female	16-24	20	4	72	4.17	75.00
6		25-34	20	4	76	4.17	79.17
7		35-44	20	4	80	Closed	-
8		45 +	100	20	100	20.83	100.00
	TOTAL		500	100	100	100.00	100.00

descriptions a scheme for chaining together the probabilities will be proposed.

Prior to doing this, however, it would be useful to examine the kind of solutions that are in principle available and to look at some of the possible problems they pose.

(1) One solution might be to add attributes independently.

(2) Another would be to add attributes conditional on one or more attributes.

(3) A third solution would be to determine each attribute simultaneously from a multi-attribute array of probabilities that have been determined using iterative proportional fitting.

Each attribute list is determined by a Monte Carlo sampling of, in each case,

(1) P(I1) P(I2) ... P(In)

(2) P(In|In-1) P(In-1|In-2) ... P(I2|I1) P(I1)

(3) P(I1,I2, ... In)

Using the first set of probabilities ignores known and measured dependencies between attributes. The third method uses all known and measured dependencies but reintroduces the "dimensionality" problem which microsimulation attempts to get around. So there is no alternative to using available probability distributions step by step in an ordered way, as in the second method, although conditional probabilities with more conditional, prior attributes can be estimated by iterative proportional fitting.

Iterative proportional fitting or IPF is a widely used technique developed several decades ago, but formalized by Fienberg (1970). Birkin (1987) has developed a general program for IPF, which can be used in probability estimation.

A typical problem which IPF could be used to solve is as follows.

Assume we have knowledge of probabilities at three different scales: small (S), medium (M) and large (L). These might be at the enumeration district scale, the local authority district scale and the national. However, we only have knowledge for one small unit and the medium and large units in which it is embedded.

At the small scale we know:

$P^S(I_a,I_b,I_c)$ = joint probability of individual attributes a, b and c for small area s

$P^S(I_d)$ = probability of an individual attribute d

We wish to add I_d to the attribute list in our population reconstruction model.

At the medium scale we know:

$P^M(I_d|I_b,I_c)$ = conditional probability of individual attribute I_d given I_b and I_c

which has been derived from

$P^M(I_b,I_c,I_d)$, the corresponding joint probability.

At the large scale we know

$P^L(I_d|I_a,I_b,I_c)$ = the conditional probability of individual attributes I_d given I_a, I_b and I_c

which is derived from the joint probability

$P^L(I_a,I_b,I_c,I_d)$.

What are the alternative ways of estimating

$P^S(I_a,I_b,I_c,I_d)$?

Using just S level information

$$P^S(I_a,I_b,I_c,I_d) = P^S(I_d)\ P^S(I_a,I_b,I_c)$$

This ignores any dependency between attributes d and attributes a, b and c. Using M level information as well

$$P^S(I_a,I_b,I_c,I_d) = P^M(I_d|I_b,I_c)\ P^S(I_a,I_b,I_c)$$

subject to

$$\sum_{a,b,c} P^S(I_a,I_b,I_c) = P^S(I_d).$$

This problem is straightforwardly solved using IPF. But what about the extra information at the large scale. This cannot be introduced directly as it may be inconsistent wwith the medium scale information. Therefore, the problem must be solved in two steps:

$$(1)\quad P^M(I_a,I_b,I_c,I_d) = P^L(I_d|I_a,I_b,I_c)\ P^M(I_a,I_b,I_c)$$

subject to

$$\sum_{a} P^M(I_a,I_b,I_c,I_d) = P^M(I_b,I_c,I_d)$$

$$(2)\quad P^S(I_a,I_b,I_c,I_d) = P^M(I_d|I_a,I_b,I_c)\ P^S(I_a,I_b,I_c)$$

subject to

$$\sum_{a,b,c} P^S(I_a, I_b, I_c, I_d) = P^S(I_d).$$

4. THE LOCATIONAL SYSTEM AND MODELS FOR MIGRATION AND MARRIAGE

4.1. The locational system

Microsimulation models normally construct attribute profiles and behaviour for samples of individuals. However, here we follow a geographical tradition (Hagerstrand, 1957; Morrill, 1965) in which whole populations are simulated. However, it is clearly infeasible to attempt the simultaneous list processing of the whole UK population (56.4 millions in 1981, 56.8 millions in 1991), but the simulation of the population of a small area should be feasible.

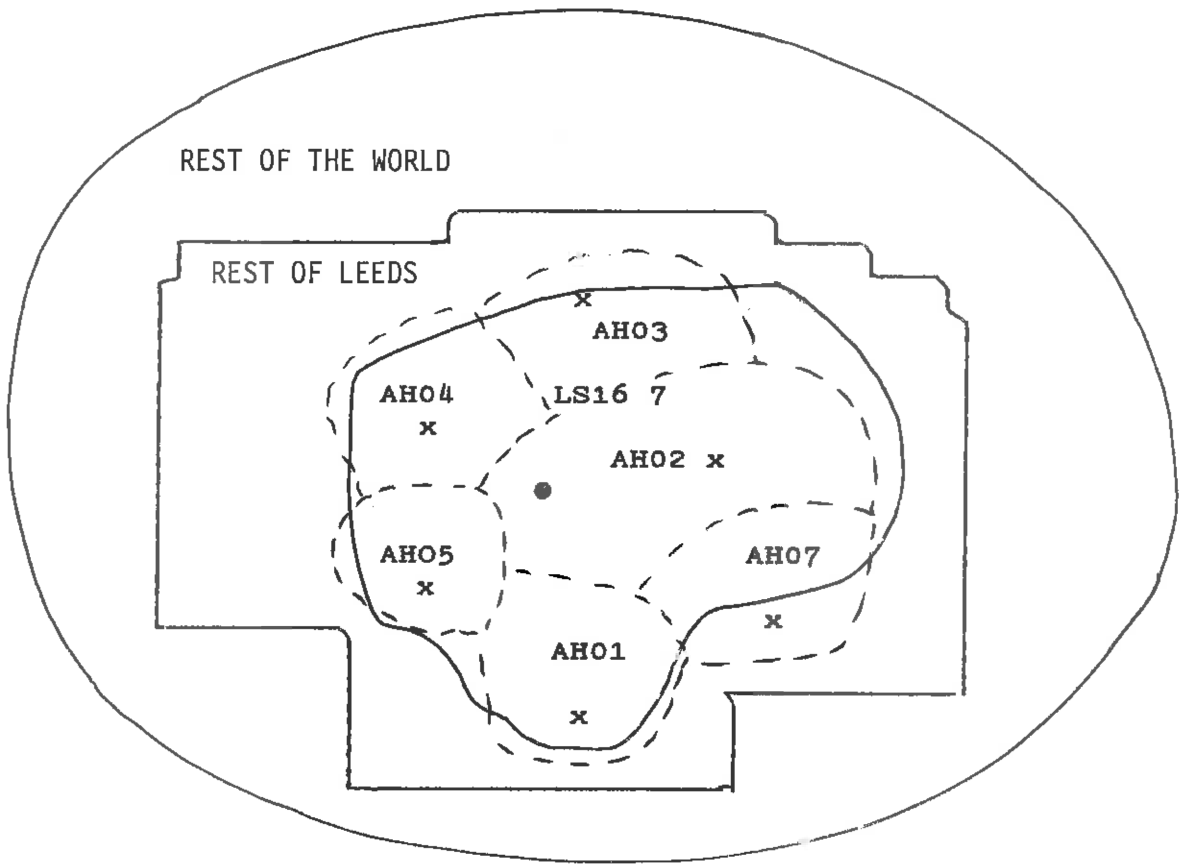
The small area to be used is the postcode sector, with circa 3,000 households and 7,500 individuals. Figure 6 shows an example (hypothetical) and the relationship of postcode sectors to enumeration districts (the smallest building blocks of the census SAS). The complex nested zones in the world outside the postcode sector will not be explicitly represented in the model, however. The idea is that once the prototype model is working for a typical small area and the relevant data banks set up, it could then be converted into a production model to run in batch mode on PIN POINT's computer to simulate as many small areas as required. There would thus be a clear distinction between the research and development stages of the project.

However, a strategy such as this assumes that the small area can be treated as an "isolated unit", which is, of course, unrealistic. Any small area is open, at varying degrees of intensity, to adjacent areas, to the rest of the local region, to the rest of the country and to the rest of the world. However, in this work we are not considering spatial interactions that are not crucial to the updating of the population: thus, we ignore interaction with workplaces, with shops, with other services, with educational establishments and so on, and concentrate on just the interactions which can permanently shift individuals and households into and out of small areas, namely migration between housing spaces.

4.2 Migration

4.2.1 Definitions

Migration, for purposes of this model, is regarded as just "a change of usual residence" by a person or by a household. Some of these migrations will be motivated by changes in employment, some for housing-related reasons and some because of life-course related events. Care will have to be taken to distinguish between the migrations of persons, either individually or together, from those of whole households. The latter, for example, will result in a housing vacancy into which another household may move, whereas the former does not often result in a vacancy in the



LS16 7 Postcode sector (PS) ● centroid of PS

AH01 Enumeration district (ED) x centroid of ED

FIGURE 6. The locational system proposed

household/housing space which is left.

4.2.2 Concepts

Note that the exact nature of migration out of or into an area must be specified: is it to be based on the transition concept (change in residence from time to time +1) or the movement concept? In the microsimulation model these two concepts lead to different treatments of out-migrants and in-migrants.

If the transition concept is adopted, this leads to

(1) the permanent removal of all out-migrants
(2) the addition of in-migrants at the end of the interval.

If the movement concept is adopted, this leads to

(1) the possibility that out-migrants might return (although this might be excluded by restricting the number of migrations to one per time interval)
(2) the addition of in-migrants during the interval and their involvement in various other processes (mortality, fertility, marriage, divorce).

This latter point raises a general issue that modellers have had difficulty in dealing with: that of ordering event sequences, which can happen in reality in many different orders, in a fixed order for computer processing.

4.2.3. Processes

Migration involves three processes:

(1) the propensity of residents in an area to migrate (the MOBILITY process);

(2) the propensity of residents in an area to select a destination within or without the area (the DESTINATION SELECTION process);

(3) the propensity of vacant housing units in the area to be occupied by households moving into the area from outside (the IN-MIGRATION process).

The MOBILITY process can be regarded as determined demographically by the individual and household attributes of the population, particularly age, sex and marital status (and life course position) together with some socioeconomic attributes and some housing attributes (particularly tenure).

The DESTINATION SELECTION process and the IN-MIGRATION process are determined in part by the nature of the housing market within the area, and the nature of the housing market in areas which constitute competing destinations. At the moment it is difficult to envisage dealing with the competing destinations in the single area model being proposed here.

4.2.4 Stocks and flows in local housing markets

Figure 7 represents stocks and flows in a local housing market. The time interval starts with the stock divided into occupied and vacant housing spaces. The vacant housing spaces constitute the housing pool at the start of the time interval. Most of the occupied housing stock continues in occupation through the year (80 to 90% on current mobility levels), but 5 to 10% will be subject to "vacations" (i.e. households moving out) that contribute to a housing pool. Other portions of the housing stock will be subject to subdivision and renovation, thus contributing housing units to the pool. A final addition to the housing pool is the set of new housing spaces built during the time interval.

This housing pool is then competed for by moving households originally resident in the area at the start of the interval, and by in-migrating households.

The housing pool can suffer minor loss, however, through demolition and amalgamation, although most of these losses will be sustained by the occupied and vacant housing stock.

4.2.5 Implications for the microsimulation model

The discussion above has crucial implications for the design of the microsimulation model. The processes that have to be incorporated are set out below (though working out the details of how this is to be done will require much further research).

(1) Each household will be exposed to a rate of moving as a whole.

(2) For each household that does not move as a whole, each household member will be exposed to a rate of moving as an individual or as part of a unit smaller than the household.

(3) The pool of moving households and fissioned part households (mainly individuals) will search for suitable housing vacancies in the local housing market. Only some will be matched to units in the small area; others will out-migrate. The successful matching of moving households with housing units in the housing pool will in part depend on a probabililstic matching of their respective attributes (size of household - size of unit; disposable household income - purchase price) and in part on the role of distance and space. We should apply a migration rate function dependent on the area's size which ensures that some movers out-migrate.

(4) The units in the housing pool not taken up by moving households resident at the start of the time interval will be available for occupation by new, in-migrating households. The characteristics of these households will be determined in part by those of the housing units and in part by the collective socioeconomic character of the area. The characteristics of these households will also be dependent on the origin area characteristics, but our single region approach excludes that influence.

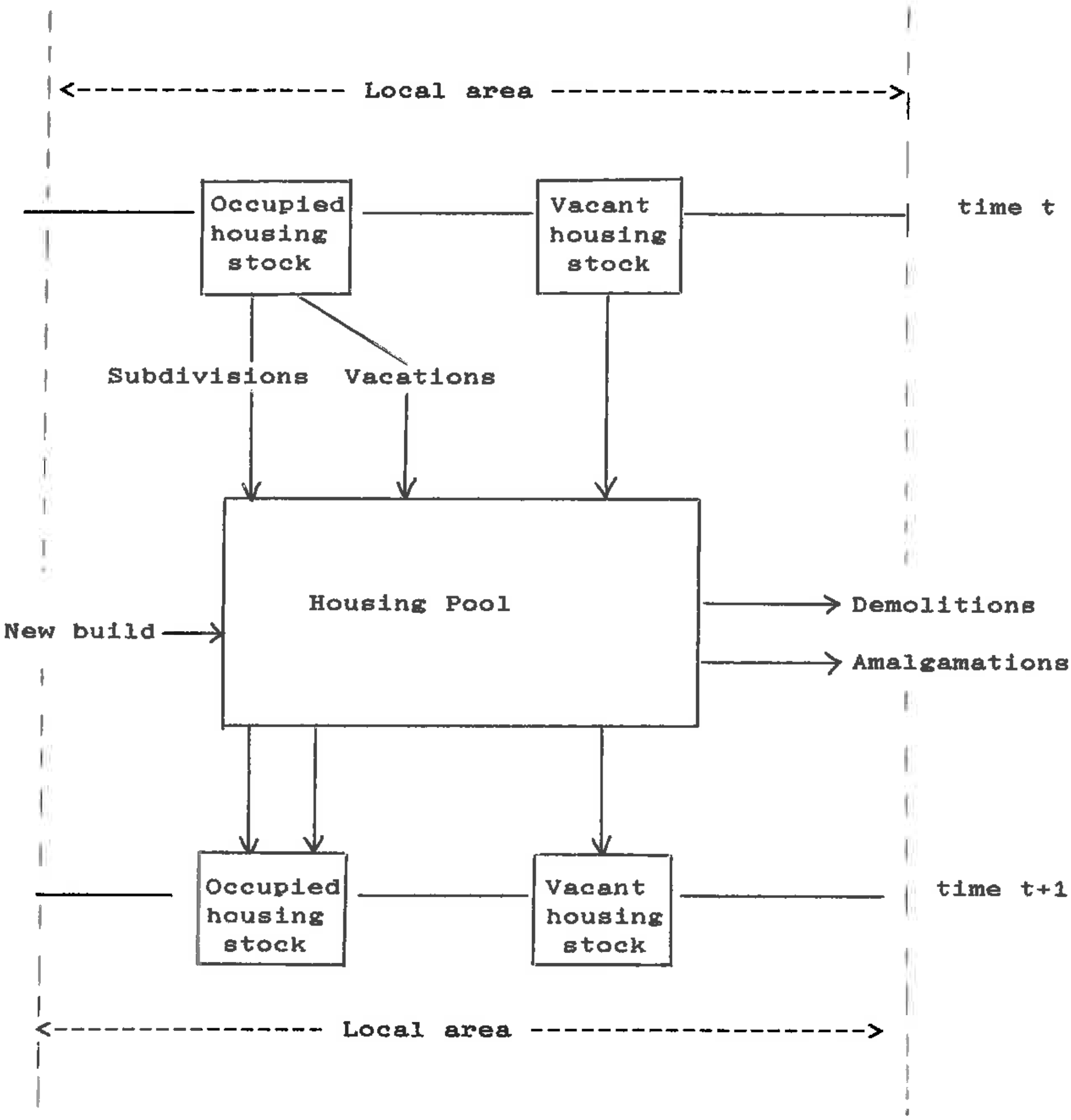


FIGURE 7. Stocks and flows in a local housing market

(5) The final consequence of this argument is that we need to incorporate in the model a list of housing units and their attributes that parallels the lists of individuals and households that the model generates. All three lists should be linked together by "pointers" (common codes).

4.3 Marital status transitions

Transitions of individuals from one marital state to another, lead to the formation of new housholds, or the transformation of old households or their dissolution, and these changes are frequently associated with migration. The essential feature of these processes is that two individuals became a couple through marriage or co-habitation, or an existing pair becomes two individuals through separation, divorce or widowhood (in this case one of the pair has died). We look first at pair formation and then at pair dissolution.

4.3.1. Pair formation

Figure 8 shows how marriage, movement and migration interact. First, the population is tested for eligibility for marriage and a pool of potential partners is created, which may include potential partners from other areas. A matching process then occurs in which potential partners are matched with either other local partners or exotic partners (partners from outside the area). The next stage involves a decision about movement: either the individual receives the partner in their existing home or the individual moves to the partner's home or both move simultaneously to a third housing unit. These decisions then intermesh with the migration process discussed above, represented by the last set of boxes which essentially involve selections of suitable destinations.

4.3.2. Pair dissolution

The equivalent set of decisions and outcomes associated with pair dissolution are setout in Figure 9. The steps are similar but simpler than in the case of marriage because no in-migration to the local area is involved.

4.3.3. Implications for the microsimulation model

The same processes would be at work in the marriage pool as previously outlined for the housing pool. The pool of local area residents eligible for marriage (i.e. likely to marry on probabilistic grounds) will search for suitable marriage partners in the local marriage market. Only some will find partners in the local area; others will find partners from outside the area. The successful matching of potential partners into marital unions will depend on the probabilistic matching of their respective attributes. Methods for carrying out such matching have been developed by Clarke and others.

It should be stressed that adoption of a microsimulation approach for marital status transitions clarifies the problem of seemingly unreasonable

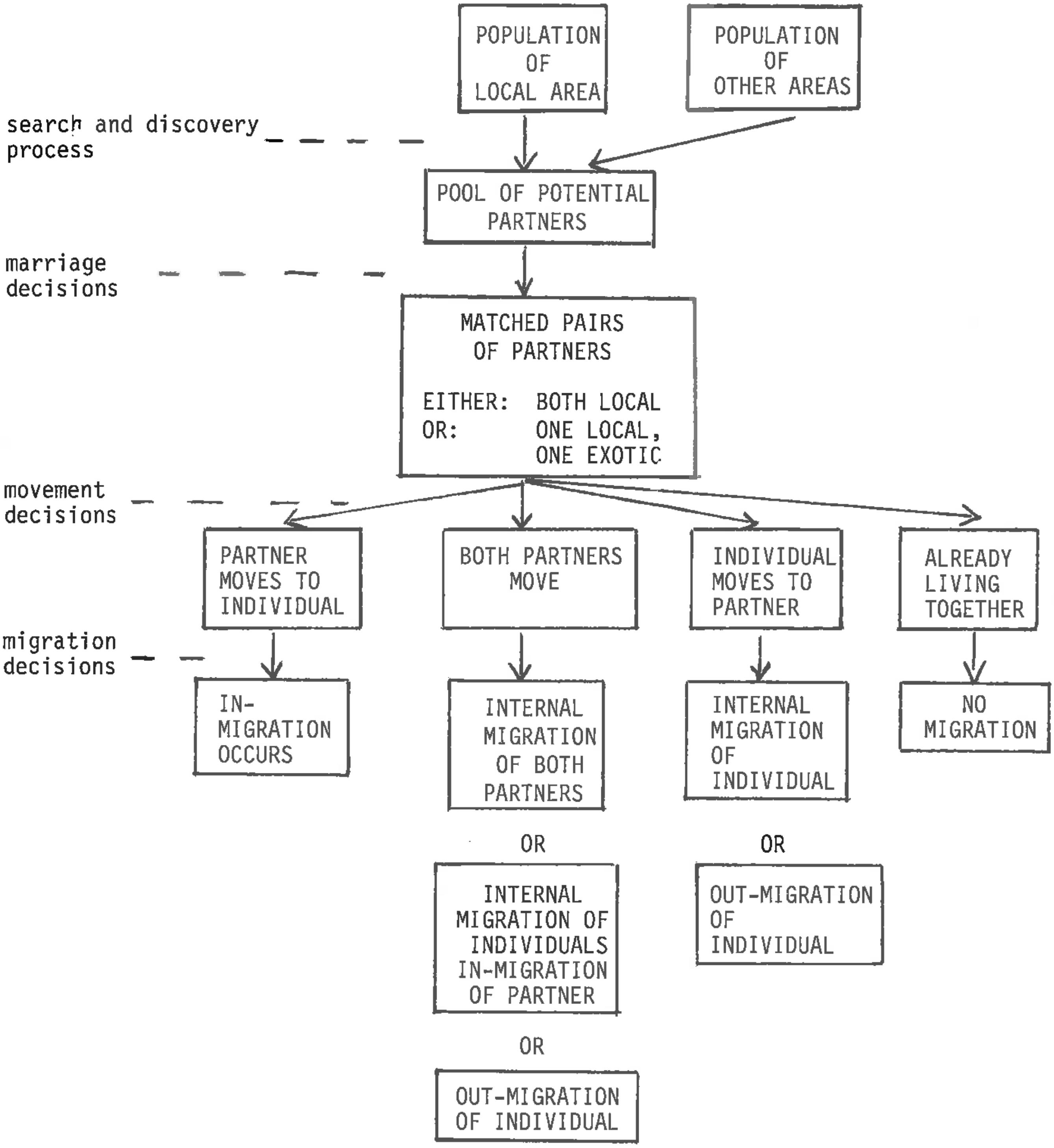


FIGURE 8. Marriage, movement and migration

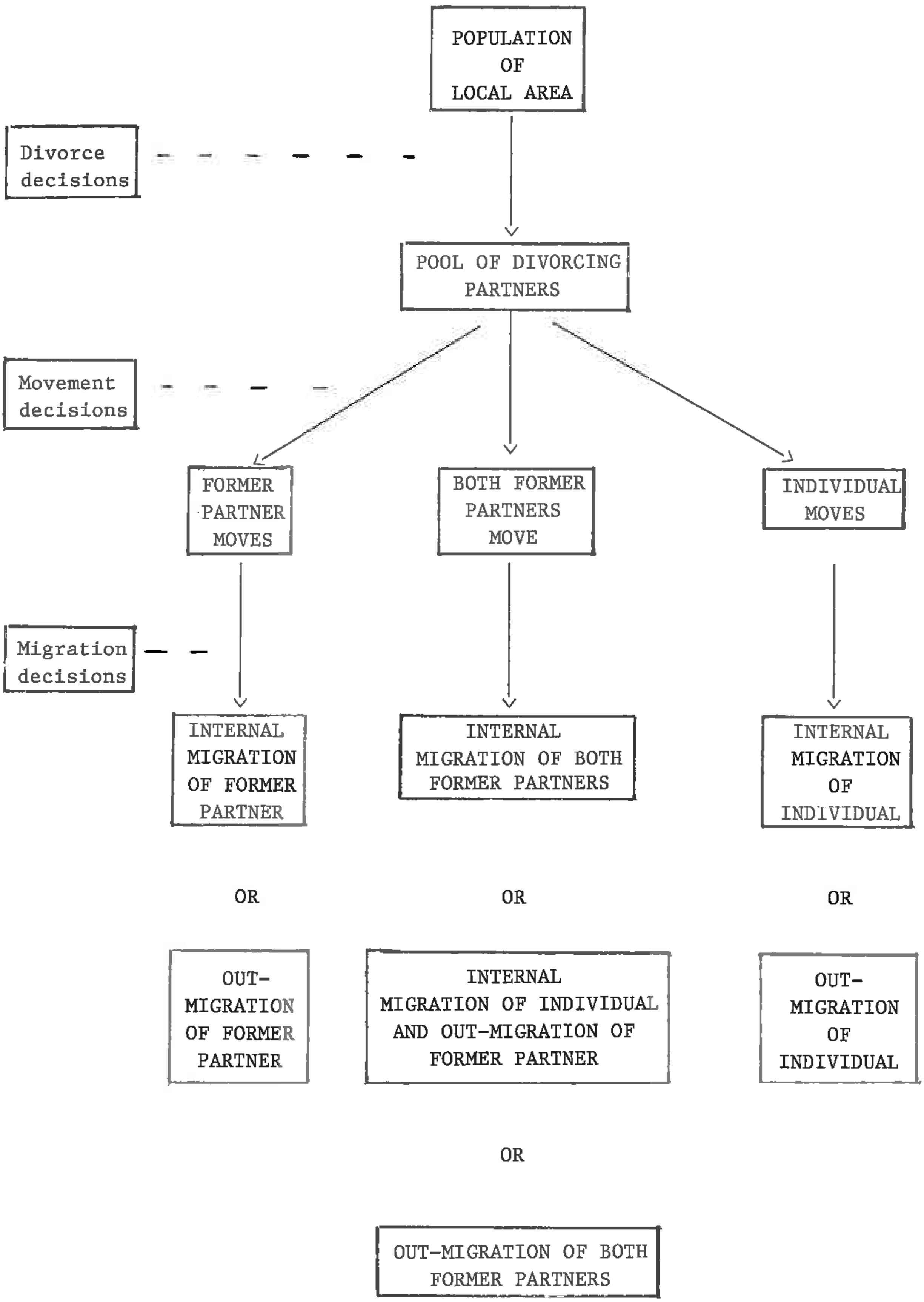


FIGURE 9. <u>Divorce, dissolution and departure</u>

transitions in macro-models. These transitions simply result from multiple events in a single time unit

e.g. single to married to widowed

which is collapsed to

single to widowed

when start point-to-finish point transition rates are used.

In the microsimulation model, such a multiple transition could occur if the population were to be "tested" for mortality after "testing" for marriage. The ordering of events in the microsimulation model is a crucial issue which has already been mentioned and to which we return later.

5. UPDATING

5.1 General considerations

In the previous section were discussed two processes that effect changes in household and individual characteristics that are crucial in the design of the model. However, there are many more transformations occurring - almost as many as there are individual attributes in the reconstructed populations of individuals, households and housing units.

The processes of transformation can be divided into two kinds:

(1) Creation and deletion processes i.e. those involved with the creation of new members of households, with the deletion of existing members, with the creation of new households and with the dissolution of old households, with the creation of new housing units and the destruction of old housing units; and

(2) processes that change the attributes of existing entities whether individual, household or housing, which persist through the unit time interval. It is useful to represent these processes in adaptations of the accounting tables used by population analysts (Figure 10). The possible starting states are classified into E, I and B categories, the exact meaning of which varies depending on the accounting entity; the possible finishing states are categorized as S, O or D. See Figure 10 for the meanings of these. For each accounting table the creation and deletion processes involve transitions EO, ED, IS, IO, ID, BS, BO and BD, whereas the transformation processes involve ES transitions. One implication of the Figure 10 framework is that new entities should be subjected the risk of O and D type events, although the risk on average will be only half that of existing entities.

Let us now consider the two sets of processes in order, with detailed consideration of each attribute where appropriate. For the moment the problem of competing risks or ordering of the transformation processes is not treated, and the following account is suggestive of approaches rather than comprehensive. Later versions of this paper will flesh out the details.

Individual transitions in a unit time interval

Starting states / Finishing states		Surviving individuals S	Out-migrants O	Deaths D	Initial stocks
Existing Individuals	E	ES	EO	ED	Starting Population
In-migrants	I	IS	IO	ID	Total In-migrants
Births	B	BS	BO	BD	Births
Final stocks		Finishing population	Total out-migrants	Total deaths	Total individuals processed

Household transitions in a unit time interval

Starting states / Finishing states		Surviving households S	Out-migrating households O	Disappearing households D	Initial stocks
Existing households	E	ES	EO	ED	Starting households
In-migrating households	I	IS	IO	ID	Total in-migrating households
New households	B	BS	BO	BD	Total new households
Final stocks		Finishing households	Total out-migrating households	Total disappearances	Total households processed

Unit (of housing) transitions in a unit time interval

Starting states / Finishing states		Surviving units E	Amalgamating O	Demolitions D	Initial stocks
Existing units	E	ES	EO	ED	Starting units
Subdivisions	I	IS	IO	ID	Total subdivisions
New construction	B	BS	BO	BD	Total new units
Final stocks		Finishing units	Total amalgamations	Total demolitions	Total units processed

FIGURE 10. A classification of t:ansformation processes

5.2 Generation of new members

5.2.1. Births (Individuals - B transitions)

Microsimulation models to date have usually adopted a female dominant fertility model. All female individuals are "tested" for maternity through application of probabilities dependent on the following key attributes which determine the fertility process: age, marital status, ethnic origin, and social class (which can be an aggregation of socioeconomic groups for present purposes). Then for each maternity event probabilities of one, two and three on births are applied (ignoring rarer multiple births perhaps).

Most fertility measurements are expressed as occurrence-exposure rates i.e.

fertility rate = fertility events/population exposed
= births/average population in time interval

Approximate probabilities are obtained through a simple transformation

maternity probability = 0.5 x (1 + survival rate) x maternity rate.

Births are now published on a ward basis each year by OPCS. Hirschfield (1986, Part VI) has demonstrated that there is significant variation across wards populations in their birth rates. Not all of this variation is due to the varying age and marital status composition of ward female populations: the age-marital status - ethnicity specific fertility rates themselves vary across wards.

How might ward level births data be used together with other fertility data for districts and the whole country to produce an annual series of maternity probabilities? This question cannot be answered fully without further detailed research. However, it is useful to outline in principle how to achieve such a series, using techniques described in Rees and Woods (1986). The description that follows, although elaborate, is illustrative of what will be required to obtain best estimates of local area probabilities in the microsimulation model. Further investigation is needed to specify whether the estimations outlined can be expressed as iterative proportional fitting problems, for which a general algorithm and program (Birkin, 1987) exists.

The finest estimates that can be made of maternity rates are for wards and we therefore need a method for using ward estimates for postcode sectors. For the moment let assume we will use the estimate for the nearest ward. What is clear is that birth figures for wards cannot be used as direct inputs to a microsimulation model for postcode sectors.

The estimation problem to be solved is as follows.

Information is available for the following fertility statistics:

B^w = births for ward w

$B^d(b,e)$ = births for district d by age of mother b and ethnicity e

$B^n(a,m)$ = births for nation n by age of mother b and marital status m

$M^n(a)$ = maternities for nation n by age of mother a

$B^n(s)$ = nations n births by number of siblings s

together with the corresponding female populations at risk

$F^w(b,e)$ = ward w population of age b and ethnic status e

$F^w(a,m)$ = ward w population of single year of age a and marital status m

$F^d(b,e)$ = district d population of age b and ethnicity e

$F^d(a,m)$ = district d population of age a and marital status m

$F^n(b,e)$ = nation n population of age b and ethnicity e

$F^n(a,m)$ = nation n population of age a and marital status m

where a are single year of age groups and b quinquennial age groups.

The target variables to be estimated are:

$P^w(a,e,m)$ = maternity probability for ward w for mothers in age a, ethnic group e and marital status m

$P(s)$ = probability that a maternity produces s infants

The solution algorithm that makes best use of the available information proceeds as follows.

(a) Birth rates for the nation are estimated.

Births are estimated first

$$B^n\ (a,e,m) = P_1{}^n\ (e|b)\ B^n(a,m) \quad a \in b \tag{1}$$

where

$$P_1{}^n\ (e|b) = B^n(b,e) / \sum_e B^n(b,e).$$

Then populations at risk are

$$F^n\ (a,e,m) = P_2{}^n(e|b)\ F^n(a,m) \tag{2}$$

where

$$P_2^n\ (e|b) = F(b,e)/ \sum_e F^n(b,e).$$

Birth rates are then computed

$$b^n(a,e,m) = B^n(a,e,m)/F^n(a,e,m) \qquad (3)$$

(b)<u>These birth rates are applied to district populations</u>:

$$B^d\ (a,e,m) = b^n\ (a,e,m)\ F^d\ (a,e,m)$$

where

$$F^d\ (a,e,m) = P^d(e|b)\ F^d(a,m) \qquad a \in b \qquad (4)$$

where

$$P^d\ (e|b) = P^d(b,e)/ \sum_e P^d(b,e)$$

subject to

$$\sum_{a \in b} \sum_m B^d(a,e,m) = B^d(b,e)$$

so that

$$B^d(a,e,m) = \left[\frac{B^d(b,e)}{\sum_{a \in b} \sum_m b^n(a,e,m)\ F^d(a,e,m)}\right]\ b^n(a,e,m)\ F^d(a,e,m) \qquad (5)$$

The district birth rates are then computed

$$b^d(a,e,m) = B^d(a,e,m)/F^d(a,e,m) \qquad (6)$$

(c) <u>These are then applied to ward population</u>

$$B^w(a,e,m) = b^d(a,e,m)/F^d(a,e,m) \qquad (7)$$

where

$$F^w(a,e,m) = P^w(e|b)\ F^w(a,m)$$

where

$$P^w(e|b) = F^w(a,e)/ \sum_e F^w(b,e)$$

subject to

$$\sum_a \sum_e \sum_m B^w(a,e,m) = B^w$$

so that the final birth estimates are

$$B^w(a,e,m) = \frac{B^w}{\sum_{a \epsilon b} \sum_e \sum_m b^d(a,e,m)\, F^w(a,e,m)} \; b^d(a,e,m)\, F^w(a,e,m) \tag{8}$$

and the birth rates

$$b^w(a,e,m) = B^w(a,e,m)/F^w(a,e,m) \tag{9}$$

(d) <u>These birth rates are converted to maternity probabilities</u>

First, maternity rates are computed

$$m^w(a,e,m) = (M^n(a)/B^n(a))\; b^w(a,e,m) \tag{10}$$

and then these are converted to probabilities

$$P^w(a,e,m) = 0.5\;(1 + s(a))\; m^w(a,e,m) \tag{11}$$

The probabilities that a maternity produces s births

$$P(s) = B^n(s)/\sum_s B^n(s) \tag{12}$$

which can be applied independently with microsimulation model.

The above equations can be applied to the base year (1980-81 or 1981), but there are problems for subsequent years in that ward populations are unknown, although crude factoring of OPCS district estimates (cf. CACI's methods) could be used.

5.2.2 <u>In-migrants</u> (Individuals-I-transitions)

The way in which these would be generated has been discussed in detail in section 4.2. They would fill vacancies in the local housing market not taken by out-migrants. The rate of out-migration must therefore be computed and applied first. Then pull factors for in-migrants must be estimated and in-migrants matched with housing units in the local area.

Three questions concerned with migration need answering.

(1) What do we use as a source of local mobility rates and destination selection rates for the base year?
(2) What sources can be used to update these migration rates?
(3) What can we use as a source of updated pull factors for in-migrants?

Tentative answers to the questions are as follows.

(1) The Small Area Statistics for the 1981 Census provide for EDs and wards the numbers of in-migrants by age and sex over the one year prior to the Census in April 1981. The MATPAC system should provide for wards the level of within ward, into ward and out of ward migration (within GB) by age and sex (unless data suppression is encountered). This information system should come on stream soon at the University of Manchester Regional Computing Centre. These statistics are for individuals, and further statistics at district, region or national level would be needed on the migration of wholly moving households.

(2) No local sources of updated mobility statistics exist, but it should prove possible to use time series of migration numbers and rates for Family Practitioner Committee Areas, which are published each quarter by OPCS. It might be possible to investigate whether changes in numbers on the electoral register reflect changes in mobility (although other factors such as entry of new electors, mortality and failure to register also cause change in the number of electors).

(3) The following sources of statistics connected with stocks and flows in the housing market should be investigated in collaboration with PIN POINT Ltd: Post Office Directory of Addresses, the Register of Electors, Housing Start and Completion Statistics, Planning Permission Statistics.

In carrying out the third stage of the model (Projection into the Future - see Figure 1), none of the above sources would be available: scenarios of new housing starts and conversions would have to be introduced into the microsimulation model.

5.2.3 New households (Households - B transitions)

These will be wholly a result of transitions occurring in the population of individuals and their attributes.

5.2.4 In-migrating households (Households - I transitions)

The creation of these households is through in-migration to vacant housing units, and has been discussed in some detail in sections 4.2 and 5.2.2.

5.2.5 New construction (Housing Units - B transitions)

The availability of housing construction statistics for local areas needs investigation.

5.2.6 Subdivision (Housing unit -I transitions)

In urban areas large old houses are converted into use as self-contained flats, particularly when properties change hands. Sources of surrogate statistics related to this process need to be investigated.

5.3 The deletion of old members

5.3.1 Deaths (Individuals - transitions to D)

Techiques for estimating local mortality probabilities similar to those discussed for births will need to be developed.

5.3.2 Out-migrants (Individuals - transitions to 0)

These transitions have been discussed above (sections 4.2, 5.2.2) in connection with migration processes and in-migration.

5.3.3 Disappearing households (Households - transitions to D)

Households will disappear wholly as a result of events to individual or decisions by them, and thus this set of transitions are simply a matter for accounting and reporting.

5.3.4 Out-migrating households (Households - transitions to 0)

Migration decisions by household members will determine how many households out-migrate in part or a whole, as discussed earlier.

5.3.5 Demolition (Housing units - transitions to D)

Of major importance in inner city areas, this process needs tracing.

5.3.6 Amalgamation (Housing units -transition to 0)

When this occurs, two housing units are merged, and one disappears for statistical purposes. This is probably a fairly minor process.

5.4 Processes that change attributes of persisting members

In the updating part of out model ("dynamics to the present"), we are dealing with individuals, households and housing units that exist at the start of a unit time interval and survive through to the end (ES transitions in Figure 10). Their attributes can change in rule-defined, in regular or in random ways. These changes are discussed for each attribute in the attribute lists developed in section 3.

(I1) Age

This simply changes with the passage of time.

(I2) Sex

This is fixed at birth and does not change thereafter.

(I3) Marital status

This changes as a result of marriage and divorce, discussed in section 4, and of the mortality of a spouse (widowhood). Changing rates of marriage, divorce and mortality will be introduced into the model.

(I5) Ethnic status

This is determined at birth by parental characteristics. Ideally, the model should recognize births to parents from different ethnic groups, creating new categories as it proceeds in an "intelligent" way.

(I6) Place of birth

All individuals born in model runs will be assigned a United Kingdom place of birth. New individuals with non-UK birthplaces will be introduced via in-migration.

(I7) Educational level

This will change rapidly in the age range 15 to 25, and a matrix of transition probabilities dependent on age and sex would need to be used. Stone (1971) has explored the construction of such matrices using 1960's data. It should be possible to update those matrices for the 1980's from the extensive body of educational statistics available. The model would again have to recognize that categories are changing (e.g. O levels are being replaced by GCSEs).

(I8) Economic activity
(I9) Employment status
(I10) Socioeconomic group
(I11) Industry

National statistics on these economic attributes are plentiful. Statistics for subnational scales are available via the National Online Manpower Information System for districts and employment office areas. Since employment changes are effected in labour markets within travel-to-work areas as a whole, these sources should provide sufficient indicators to provide transition probabilities for local areas within these TTWAs. However, further investigation is clearly indicated.

(I12) Earned income

Here we will have to rely almost exclusively on national surveys (New Earnings Survey, Family Expenditure Survey, Labour Force Survey). Earned incomes will be generated from probability distributions conditional on occupation, industry category, age and sex. A method for generating individual incomes has been developed in a seperate research project, details of which can be found in Clarke and Clarke (1987).

(H1) Type of household

(H2) Size of household

These will be the result of various individual actions or events, and just need careful accounting and reporting.

(H3) Earned income

This will be sum of the earned incomes of individuals in the household.

(H4) Benefits

Complex sets of rules govern the provision of benefits depending on the attributes of individuals. The principal, non-means tested benefits can be computed by rule (e.g. child benefit, old age pension) but others (disability allowance, unemployment benefit) will have to be computed probabilistically.

(H5) Unearned income

This will be the sum of unearned incomes of individuals in the household.

(H6) Taxation

The same comments apply as in the previous paragraph.

(H7) Net household income

This will be the result of adding (H3), (H4) and (H5) and subtracting (H6).

(H8) Number of cars

Survey figures (e.g. General Household Survey) provide information on the ways in which the distribution of households across car categories is changing, but from these cross-sectional figures we need to estimate transition rates. At the very minimum the mid-point between possible extremes of independence and minimum transition could be chosen - see Table 5 for an illustration.

(H9) Tenure

This will change principally through migration (but see U2). National Surveys may provide some data on previous tenure of households interviewed. Otherwise estimates can be made using updated vectors (as H8).

(H10) Rooms

Changes in the number of rooms households occupy will principally be effected through migration (but see U3).

(H11) Amenities

Changes will probably occur mainly through improvement rather than migration (see U4).

(U1) Occupance

Changes in this variable would be a result of the household-housing unit matching process in the microsimulation model.

(U2) Tenure

TABLE 5. An illustrative calculation of transition probabilities from known marginals

5.1 Census figures

Number of cars in household	Census 1971 %	Census 1981 %
0	49	39
1	42	46
2 or more	9	15
Totals	100	100

5.2 Minimum transitions

t+10		Census 1981			Totals
t		0	1	2+	
Census	0	39	10	0	49
1971	1	0	36	6	42
	2+	0	0	9	9
Totals		39	46	15	100

5.3 Transitions under hypothesis of independence

t+10		Census 1981			Totals
t		0	1	2+	
Census	0	19	23	7	49
1971	1	16	19	7	42
	2+	4	4	1	9
Totals		39	46	15	100

5.4 Mid-point between extremes

t+10		Census 1981			Totals
t		0	1	2+	
Census	0	29	16.5	3.5	49
1971	1	8	27.5	6.5	42
	2+	2	2	5	9
Totals		39	46	15	100

Changes will occur as a result of the sale of publicly owned property to sitting tenants, changes from leasehold to freehold and from owned to rented on conversion. The housing statistics will give some guidance here.

(U3) Rooms

As a result of housing improvement, existing units could have their numbers of rooms increased. The housing statistics may give some guidance here.

(U4) Amenities

Although the 1970s and 1980s have seen a substantial diminution of the number of housing units lacking exclusive use of all amenities, certain tenure categories and particular areas still show a high incidence of amenity deprivation. For example, 27.8% of households renting private furnished accommodation lacked exclusive use of an inside WC in 1981. It is important therefore to monitor improvements in amenity provision in local areas.

6. THE WAY FORWARD

In this paper a model for updating populations of individuals and households in small areas has been sketched out. At the moment this model exists a series of steps, techniques and ideas all of which need further investigation, elaboration, integration and ordering.

6.1 Summary of the model so far

The ordering of events in the model is particularly important. Figures 11, 12 and 13 set out a suggested ordering. The population reconstruction phase begins with the creation of households and new members (Figure 11). For each member of the household demographic, educational and economic attributes are determined. Then the household loop continues with a recording of the type and age, a generation of earned income, benefits, taxation and disposable income summing the incomes of individual and applying benefit rules based principally on family composition. The number of cars per household and other consumer durable possessions can then be estimated on the basis of household income estimates followed by the nature of the housing unit occupied by the household (tenure, rooms, amenities). The final steps in the household loop are then to check all constraints, close any full categories, recompute any probabilities and create a housing unit record. A new head and household can then be created.

In the microsimulation phase of the model the same nested household and individual loop structure is used: members are subject to mortality, fertility, marriage, divorce and migration risks in that order (Figure 12). If the relevant events occur side loops are activated which result in the subtraction or addition of individuals, the creation or deletion of households and changes in some of their attributes. Where new members have been created or new households have been formed these must be processed once the loop for existing members and households have been completed.

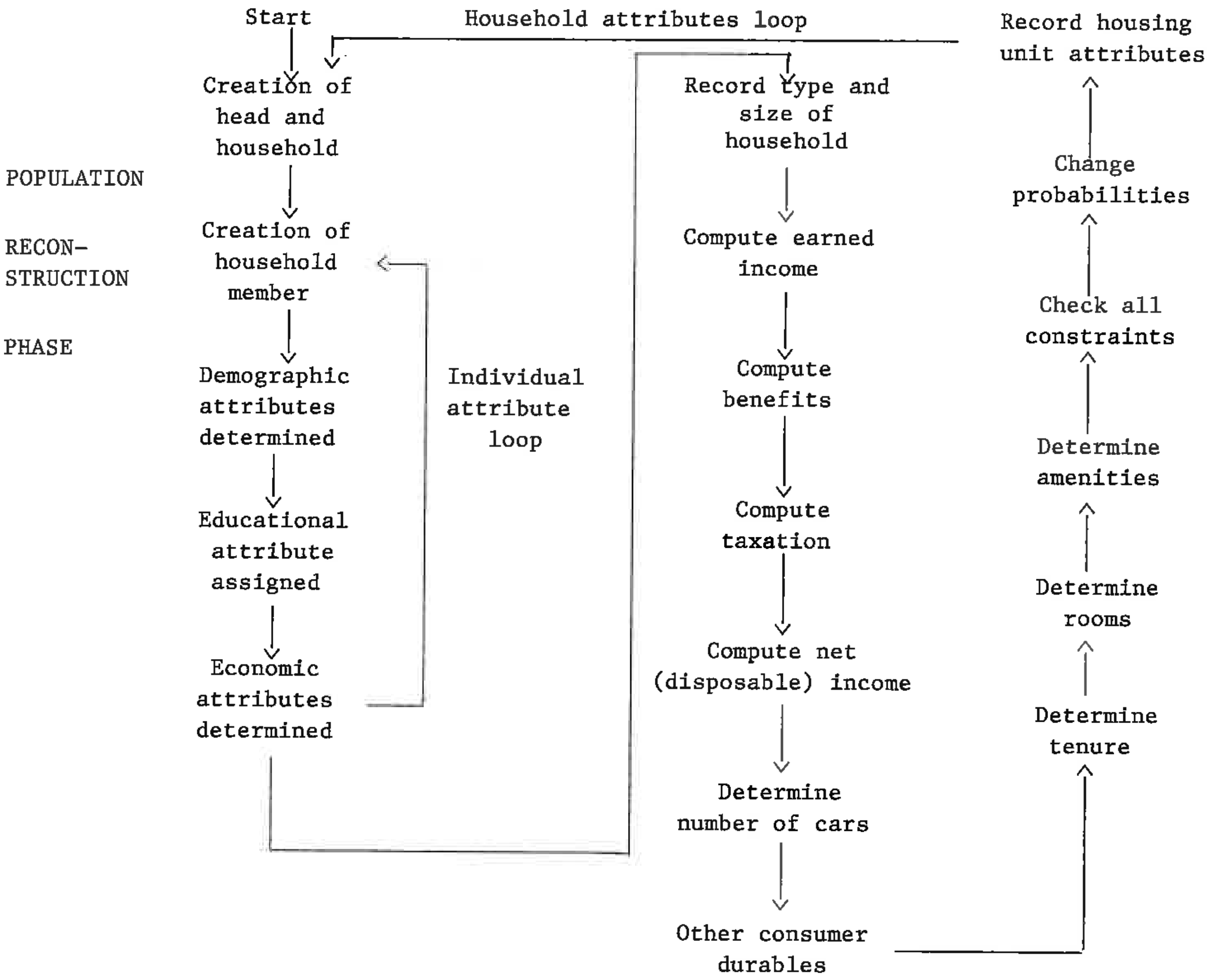


FIGURE 11. Summary flow diagram of population reconstruction phase of the model

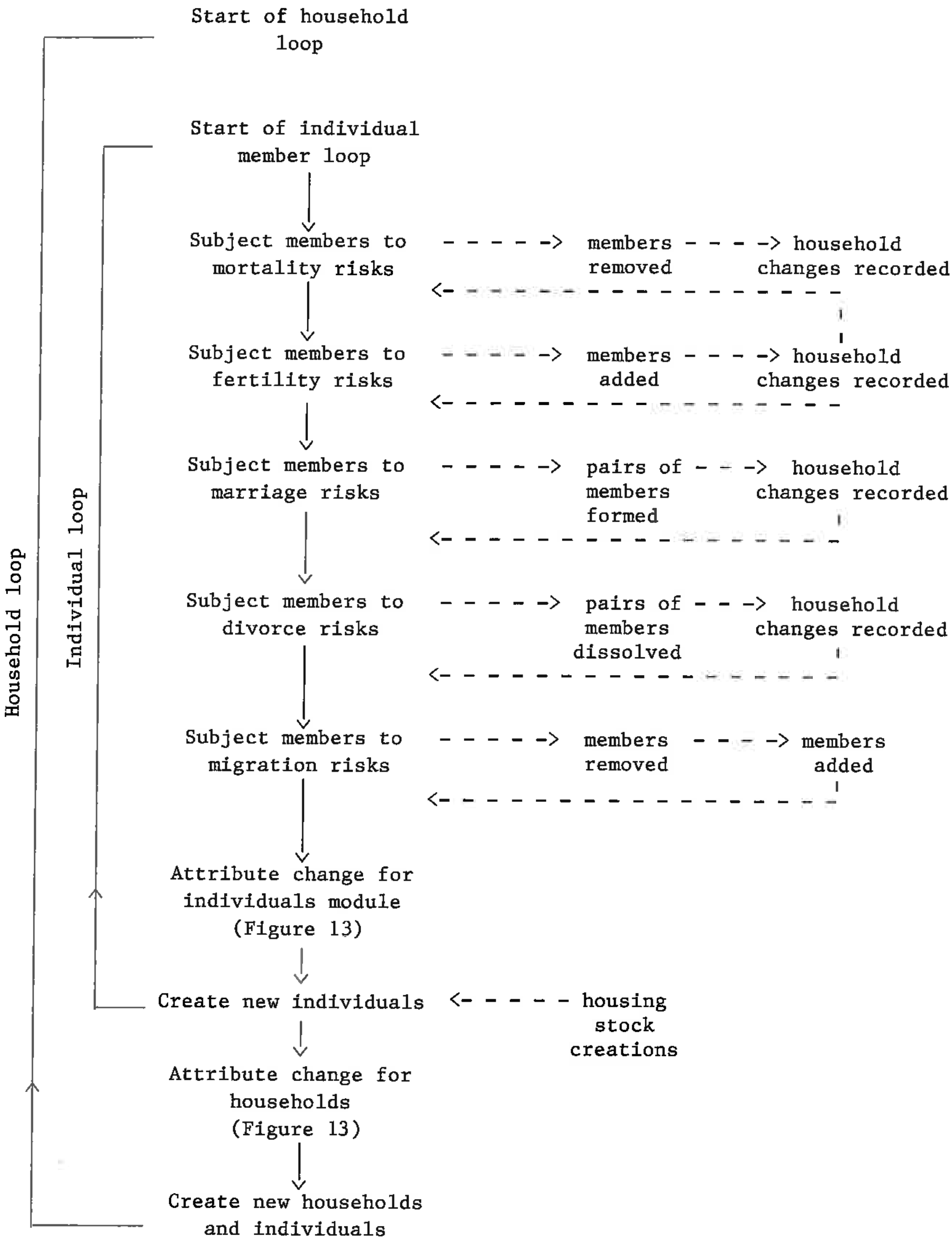


FIGURE 12. The creation and deletion phase of the model

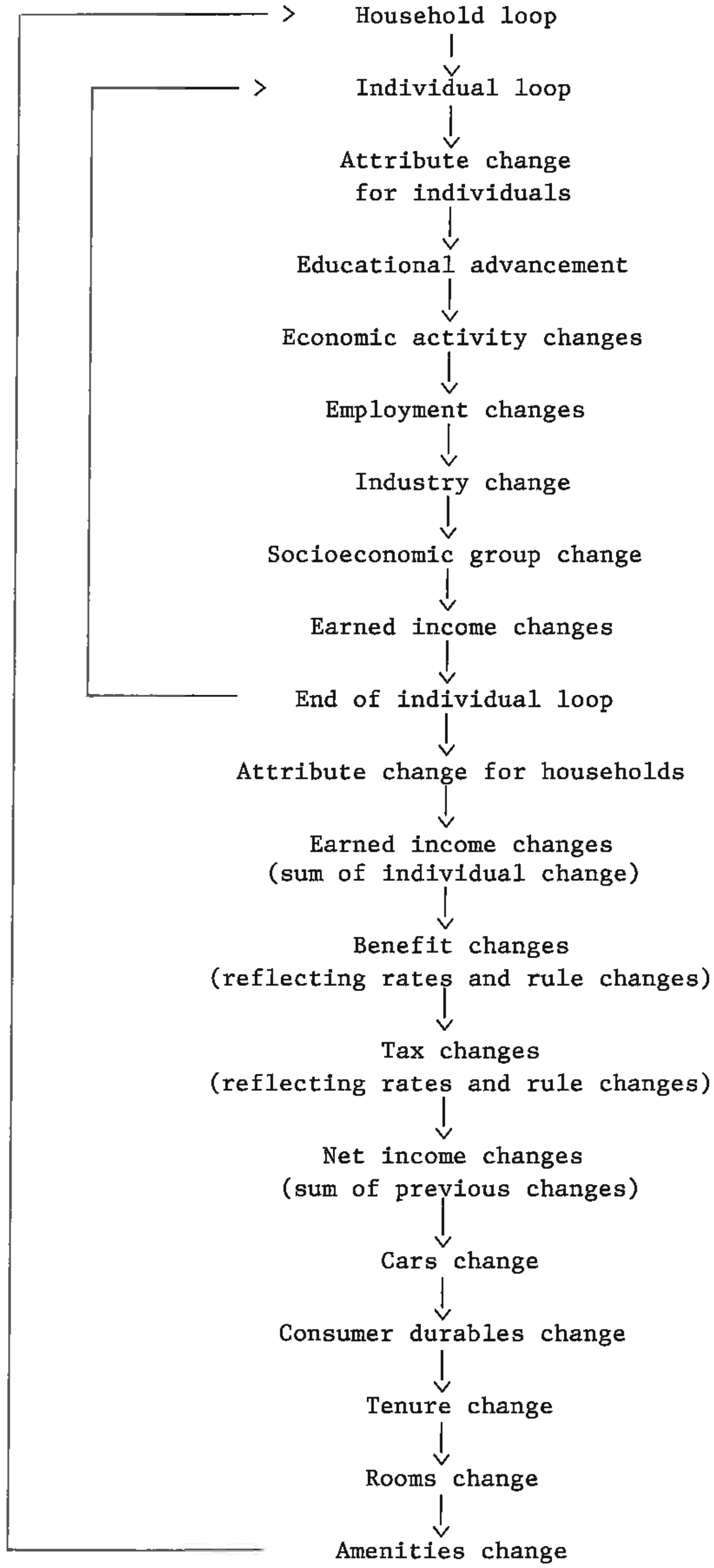


FIGURE 13. The socioeconomic change phase of the model

The final phase of the model subjects households and individuals to socioeconomic change (Figure 13).

6.2 Ethical issues

A final comment is appropriate on the ethical issues that the model raises for individual data privacy, because the data inputs to the model are derived from the returns of individuals and households to administrative bodies (OPCS, Electoral Registration Office, the Post Office) and there is concern that such information be treated in complete confidence (Campbell and Connor, 1986). Two data protection principles apply:

(1) data about individuals should not be passed on to third parties without the knowledge and consent of the data subject; and

(2) the data should only be used for the purposes for which it was gathered.

Given that this model attempts to simulate the lives of the <u>entire</u> population of a small area we must be very careful

to

(1) never include names or addresses in any file of data that we use.

(2) never include information on simulated individuals or households in published reports.

(3) only produce tables of statistics in such a form that simulated individuals or households cannot be reconstructed.

In this way there should be no opportunity to confuse a simulated individual or household with a real one.

REFERENCES

OPCS = Office of Population Censuses and Surveys
COI = Central Office of Information

BIRKIN, M. (1987) Iterative proportional fitting (IPF): theory, method, and examples. Computer Manual 26, School of Geography, University of Leeds.

BIRKIN, M. and CLARKE, M. (1986) RUIN: a really useful information system. Paper presented at the British Section, Regional Science Association, Annual Meeting, School of Advanced Urban Studies, University of Bristol, September 1-4, 1986.

CAMPBELL, D. and CONNOR, S. (1986) on the record: computers, surveillance and privacy. Michael Joseph, London.

CHANDOR, A. with J. GRAHAM and R. WILLIAMSON (1985) The Penguin dictionary of computers. Third Edition. Penguin Books, Harmondsworth, Middlesex.

CLARKE, G. and CLARKE, M. (1987) A method for generating individual incomes at the small area level. Working Paper, School of Geography, University of Leeds.

CLARKE, M. (1986) Demographic Processes and household dynamics: a microsimulation approach in WOODS R. and REES, P. (eds) Population Structures and Models. Allen and Unwin, London.

CLARKE,M. and HOLM,E. (1987). Microsimulation methods in spatial analysis and planning. Geografiska Annaler, forthcoming.

DULEY, C. (1987) Updating census-based household and population information for inter-censal years. Research Support Group Paper, School of Geography, University of Leeds.

FIENBERG, S. (1970) An iterative procedure for estimation in contingency tables. Annals of Mathematical Statistics, 41, 907-917.

HAGERSTRAND, T. (1957). Migration and area: survey of a sample of Swedish migration fields, and hypothetical considerations on their genesis. In D. HANNERBERG (ed.) Migration in Sweden: a symposium. Lund Studies in Geography Series B. Human Geography No. 13, Lund, Sweden. Pp. 27-158.

HIRSCHFIELD, A. (1986) Urban deprivation: its causes, manifestations and inter-relationships with urban social structure and health status. An investigation with specific reference to Leeds Metropolitan District. Ph.D. thesis, School of Geography, University of Leeds.

MORRILL, L. (1965) "The Negro ghetto: problems and alternatives" Geographical Review 55, 2, 339-361.

OPSC (1984) Census 1981. Household and family composition, England and Wales. CEN 81. HFC. HMSO, London.

OPCS and COI (1986). Households. Census Guide 4, HMSO, London.

REES, P.H. and Woods, R.I. (1986). Demographic estimation. Chapter 12 in WOODS, R.I. and Rees, P.H. (eds.) Population structures and models. George Allen & Unwin, London. Pp 301-343.

STONE, R. (1971) Demographic accounting and model building. OECD, Paris.