WORKING PAPER 515

A MICROSIMULATION MODEL FOR UPDATING HOUSEHOLDS IN SMALL AREAS BETWEEN CENSUSES

CHRISTOPHER DULEY, PHILIP REES AND MARTIN CLARKE

School of Geography
University of Leeds
Leeds LS2 9JT

October, 1988

Paper presented at the Workshop on "Multistate demography: measurement, analysis, forecasting", October 31-November 4, 1988, Kerkebosch Castle, Zeist, The Netherlands, organized by the Netherlands Interuniversity Demographic Institute and the University of Utrecht.

CONTENTS

List of table
List of figures
Acknowledgments
Abstract

1. INTRODUCTION

1.1 The multistate population modelling problem
1.2 The microsimulation solution
1.3 Microsimulation applications
1.4 Aims of the current study

2. POPULATION RECONSTRUCTION: STRUCTURE AND METHODS

2.1 The micro units and the small area populations
2.2 The attributes
2.3 The structured creation of household members
2.4 The probability equations
2.5 Probability estimation techniques

3. POPULATION RECONSTRUCTION: EXAMPLES AND TESTS

3.1 Simulated individuals
3.2 New crosstabulations
3.3 Tests of the simulations

4. UPDATING: VITAL EVENTS

4.1 The locational system
4.2 The estimation of mortality probabilities
4.3 The estimation of maternity probabilities
4.4 Illustrative results

5. UPDATING: FURTHER PROCESSES

5.1 Marriage and cohabitation; divorce and dehabitation
5.2 Migration
5.3 Socioeconomic transitions

6. CONCLUSIONS

REFERENCES

LIST OF TABLES

1. The list of individual attributes.
2. The list of household attributes.
3. The list of housing unit attributes.
4. An illustration of the procedure to extend attribute categories.
5. The estimation of the age combinations of married couples.
6. Thirty three typical tykes.
7. New crosstabulations: examples for postcode sector LS6 1.
8. A comparison of observed and predicted populations in two Leeds postcode sectors: Z scores.
9. Fertility and population statistics used in the estimation of ward maternity probabilities.

LIST OF FIGURES

1. The general structure of UPDATE.
2. The household classification.
3. Household, family and individual loops in the reconstruction process.
4. Stages in the generation of dependent children.
5. Birth rates for mothers in Headingley ward, Leeds.
6. Stages in the simulation of cohabitation and marriage.

ACKNOWLEDGMENTS

The authors would like to thank the Economic and Social Research Council for their support of the research via a CASS studentship award to the first author, and to thank Pin Point Ltd for their additional support and collaboration. We trust their patience will be rewarded. The second author would like to thank the Nuffield Foundation for support in the academic year, this project providing him with insight into the household structure of the British population.

ABSTRACT

The paper provides an interim report on the construction of a general microsimulation model designed to update the numbers and attributes of small area populations between decennial censuses. The procedures for reconstructing the base population as lists of individuals, households and housing units are described in detail. The methods for estimating small area mortality and maternity probabilities for use in the death and birth stages of the simulation are outlined. Approaches to the incorporation of the further processes of marriage, cohabitation, divorce, dehabitation, migration and socieconomic change in the model are sketched and will be the subject of future reports.

1. INTRODUCTION

1.1 The multistate population modelling problem

There is, in social scientific endeavours, a constant tension between wishing to deal with human behaviour and man-environment relations in all their complexity and wishing to make firm predictions about such matters for large numbers of people. On the one hand the anthropologist makes a detailed study of the mores of a small community; on the other hand the demographer projects selected characteristics of the whole population of a nation. We know that the first approach may suffer from the sin of particularity in which the results are applicable only to that one community. The second approach imposes an artificial homogeneity on a population: the diverse character of the individuals that make it up will profoundly affect the key variables of survival, fertility and migration employed by the demographer.

Multistate methods represent one approach to the heterogeneity problem in macro analysis. The population is divided into sub-groups and the behaviour of the sub-groups is measured and then represented in the demographic model. So, for example, the population is almost always broken down into sub-groups distinguished by age and sex, and quite often by marital status. Less frequently region of residence is introduced as a distinguishing feature. Among other dimensions of interest have been labour force status, ethnicity, occupation and income status. What happens when we disaggregate the population is that the number of variables involved in say, a demographic projection model, begins to grow. As the number of population sub-groups grow we begin to worry about whether the rates of transition from one state to another can be reliably measured and projected.

In general, if one recognizes n sub-groups in the population, then over one time interval there will be n^2 transfers between those sub-groups. The square of n grows alarmingly as n is increased. But, you object, not all transitions occur and need representation in the model. Persons only move from one age to the next in a time interval. Persons do not usually change sex or ethnic group. More than half the moves between marital statuses are disallowed (the married can never re-enter the never-married state).

However, for many other characteristics of the population, including social class and spatial location, transitions amongst all states are possible, and very often of great interest. As we incorporate more and more interesting characteristics into our model that research tells us affect the dynamic behaviour over time of populations, then the number of model variables or elements, E, increases as a function of

$$E = \prod_{a=1}^{z} N_a$$

where N_a is the number of categories or classes recognized for attribute a, where there are z such attributes incorporated in the model, and the terms are multiplied together.

Many solutions to the dimensionality problem have been proposed, most of which involve some form of aggregation of the transitions allowed in the model or decoupling of the system being studied into several component subsystems (Rogers 1976). For example, in multiregional projection models, the analyst may retain a large number of regions but settle for 5 year age groups and five year time intervals as a consequence. Frequently, however, the potential user of such projections requires single years of age data (particularly for educational planning), and projected populations on an annual basis. The solution is then to limit the number of regions involved, which usually means that the simultaneous, interdependent projection of a set of small region populations is not possible. A reduced model (bi-regional or tri-regional) must be applied independently to each area in turn.

Alternatively, inter-state transitions in the model are collapsed in one way or another. For example, in the model used to produce the English subnational population projections (Martin Voorhees and Bates 1981; reviewwed in Rees and Willekens, 1986), rates of total out-migration from each one of the 108 subnational areas are estimated (using model migration schedules for 12 clusters of areas for the 2 sexes) for single years of age to 75 years and over, but inter-area migration rates are applied for only three age groups (0-16 with 30-59; 17-29; 60 and over), and model migration schedules used again to disaggregate destination totals. Representing all inter-area migrations by sex and age would involve using

2 x 76 x 108 x 108 = 1,772,928 variables

whereas the reduced representation involves as input variables only

(2 x 76 x 12)	origin variables
+ (2 x 3 x 108 x 108)	distribution variables
+ (2 x 76 x 12)	destination variables
= 73,632 variables.	

It is useful to formalize these observations about the nature of the compromises involved in constructing a large multistate model.

Assume that we are interested in projecting the population classified by n dimensions from time 1 to time 2: that is,

$K(a_1, b_1, \ldots, n_1)$

becomes

$K(a_2, b_2, \ldots, n_2)$

where K stands for population, a,b, ... , n for the various characteristics of interest, and 1 and 2 are the consecutive points in time.

A completely specified model would involve the following:

$$K(a_2, b_2, \ldots, n_2) = P(a_2, b_2, \ldots, n_2 | a_1, b_1, \ldots, n_1) K(a_1, b_1, \ldots, n_1) \qquad (1)$$

where $P(a_2, b_2, \ldots, n_2 | a_1, b_1, \ldots, n_1)$ is the conditional probability that the population is in states a_2, b_2, ... ,n_2 at time 2 given that it was in states a_1, b_1, ... ,n_1 at time 1. The conditional probability is derived by dividing the joint probability distribution by the time 1 marginal probabilities. So, the essential task in projection is to develop a suitable model for the joint probability distribution

$P(a_2, b_2, \ldots, n_2, a_1, b_1, \ldots, n_1)$.

At one extreme lies the hypothesis of independence:

$$P(a_2, b_2, \ldots, n_2, a_1, b_1, \ldots, n_1) = P(a_2)\, P(b_2) \ldots P(n_2) \times P(a_1, b_1, \ldots, n_1) \qquad (2).$$

More sensible would be to use conditional probabilities for each dimension:

$P(a_2, b_2, \ldots, n_2, a_1, b_1, \ldots, n_1)$

$$= P(a_2|a_1)\ P(b_2|b_1)\ \ldots\ P(n_2|n_1)$$

$$\times\ P(a_1, b_1, \ldots, n_1) \qquad (3)$$

but this would capture little of the interdependency between attributes. A more complex model would be

$$P(a_2, b_2, \ldots, n_2, a_1, b_1, \ldots, n_1)$$

$$= P(a_2|a_1)\ P(b_2|a_1, b_1)\ \ldots$$

$$P(n_2|k_1, l_1, n_1)$$

$$\times\ P(a_1, b_1, \ldots, n_1) \qquad (4)$$

where successively more complex conditional probabilities would be used.

The art of multistate modelling resides in the selection of chains of conditional probabilities that capture the most significant interdependencies between population characteristics, and that can be reasonably measured from available data. Although many multiregional projection applications (e.g. the Migration and Settlement studies of the International Institute for Applied Systems Analysis, Rogers and Willekens 1986) appear at first sight to be full state-space models (as equation 2), on deeper examination it will be found that many of the input variables had to be estimated from partial data using aggregated conditional probabilities (as in equation 4).

1.2 The microsimulation solution

Microsimulation models represent one strategy for specifying and implementing the chains of probabilities in complex situations. The successive probability distributions are sampled through random selection and the characteristics are assigned to synthetic individuals. These individuals will eventually be assigned all the characteristics considered important for the problem being studied, and they can be changed as time proceeds.

A simple illustration of what goes on at the heart of a microsimulation model will help at this stage in fixing ideas. Assume that we wished to assign ethnic group membership to each household head in the model. From available small area and other statistics, we work out that the probability of being British in ethnic origin is 0.87 and of being non-British is 0.13. These probabilities are cumulated as 0.87 probability of being British and 1.00 of being British or non-British. A random number is then generated in the range 0.01 to 1.00, say 0.89. This falls within the range 0.88 to 1.00, with the consequence that the characteristic non-British is assigned to the household head being simulated. The next random number might be 0.17, and so the next head of household would be placed in the British group.

Tackling the multistate modelling problem at the level of the individual actor in the system has a number of advantages.

(1) It is easier to think about inter-dependencies of attributes at this scale, and much easier to link individuals together in groups, such as families or households. Modelling the temporal evolution of households using aggregate methods is highly artificial and highly dependent on special purpose surveys.

(2) The information storage requirements of such microsimulation models are much smaller for large populations than are the requirements of explicit multistate aggregate models. If there are M individuals in the population and n characteristics being represented, the basic information stored will be Mn elements plus some function of the sum of

N_a input elements rather than the E elements (the product multiplication of the number of categories of each attribute included).

(3) In microsimulation models, eclecticism about data sources is common: probabilities are used from national, regional and local sources as available. This approach is also possible in aggregate models, but is not often adopted.

(4) Finally, because the output of microsimulation models is a sample of individuals, flexible tabulation of characteristics is very easy. This flexibility is also possible in aggregate models, but is rarely achieved.

1.3 Microsimulation applications

Clarke and Holm (1987) have recently provided a detailed review of microsimulation methods in spatial analysis and planning. Here we summarize very briefly some of their main findings. Microsimulation methods have been used over many decades by researchers in Sweden and the USA, and more recently over the past decade in Britain. The variety of applications is very wide and includes dynamic simulation of income (Orcutt *et al* 1976), housing system analysis (Kain and Apgar, 1985), an income distribution and jobs simulation (Betson, Greenburg and Kasten, 1980), longitudinal modelling of individuals and households (Galler and Wager 1986), innovation diffusion (Hagerstrand 1952, 1967), health care planning (Clarke and Spowage 1985), household dynamics (Giesbrecht and Field 1969; Clarke 1986), population activity systems (Holm 1984; Holm and Oberg 1984; Holm, Makila and Oberg 1985) and combined macro-micro models (Clarke and Wilson 1986).

Clarke and Holm (1987) also identify several potential novel applications of microsimulation methods, one of which is the development of information systems for urban and regional systems. Such a system, called SYNTHESIS (Synthetic spatial information system) has been constructed by Birkin and Clarke (1988) for the City of Leeds that links, at the small area scale, census variables, income/expenditure characteristics and shopping activity patterns for one point in time. The model described in this paper can be placed in the same category, but the focus is on the development of individual and household numbers and characterictics over time, rather than on detailed interactions in a spatial system. Simultaneous attention to detailed spatial activity patterns and detailed population dynamics is a future area for microsimulation model development.

1.4 Aims of the current study

The problem which the model described in this paper addresses is that of information decay between national censuses, normally carried out at ten year intervals around the beginning of each decade. Currently in Britain, both public and private sector decision makers continue to use the 1981 Census data in situations where geographical detail is essential, for example, in the targetting of area based policies of urban revival or in the targetting of direct mail marketing campaigns. However, such small area census data dates rapidly both as a result of national social and economic trends, and through local processes of housing demolition, decay, redevelopment and new construction, together with the associated processes of population mobility and housing turnover.

The broad aim of the UPDATE model, described in subsequent sections of the paper, is to improve on the census geographic data base by updating it. A limited amount of updating is carried out for aggregate population counts by age and sex by the Office of Population Censuses and Surveys (OPCS) in their district estimates (e.g. OPCS 1988) or by marketing analysis firms such as CACI at the ward level. UPDATE aims to reconstruct a base population measured at a previous census (1981) in small areas as a list of simulated individuals and households with a variety of interesting attributes, and then to change those individuals and households over time to the present using fresh information about demographic events and socioeconomic trends to alter both the numbers and characteristics of the simulated population. The updating process will be taken forward from the present (usually last year given information lags) to the next census both to provide a population projection and to make possible a test of methods. 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 small area and other information as possible. Section 2 of the paper outlines the methods used in the reconstruction and section 3 exemplifies the techniques and outputs of this model phase. The second phase of UPDATE involving changing the numbers and characteristics of individuals in each small area population by applying a series of demographic and socioeconomic processes. Section 4 of the paper discusses the methods used to handle birth and death, with illustration of sample results. Section 5 describes the ideas that underpin the updating processes upon which work is still proceeding. Section 5.1 outlines the submodel which has been developed to handle marital union and dissolution, and, importantly,, its informal equivalent, cohabitation and dehabitation. Section 5.2 describes the planned approach to introducing migration into the model and section 5.3 briefly deals with socioeconomic transitions. The paper thus represents an interim report on the construction of a complex microsimulation model, previously blueprinted in Rees, Clarke and Duley (1987).

2. POPULATION RECONSTRUCTION: STRUCTURE AND METHODS

2.1 The micro units and the small area populations

In aggregate multistate models, the implied micro unit has normally been the individual person. Connection to household units has been via headship rates and the analysis of trends in such rates (e.g. Corner 1987). However, changes in household numbers and composition arise much more naturally from events that happen to individuals (marriage, divorce, birth, death, migration). In the UPDATE microsimulation model simultaneous tracking of individuals and the households they belong to is carried out, and what are complex household transitions that could arise from a variety of individual events at the aggregate level simply involve proper accounting in the microsimulation model. The UPDATE model, however, recognizes a group intermediate between the individual and the household, namely the family and adopts the 17 household type classification (types A through Q) of the British census. Figure 2 sets out the definitions of these types and gives the distribution of population across the types by both households and persons. In terms of frequency, the most common type of household is the married couple without children (before or after childraising) making up 24% of households in England and Wales in 1981, followed closely by married couples with dependent children only (23%) and persons living alone (22%). These three types make up 69% of households in 1981, with 8% of households being some kind of lone parent household, 17% other forms of married couple household (8% with non-dependent chidren only), 5% in non-family households with 2 or more persons and under 1% in multifamily households. Note that when the distribution of persons across household types is considered the importance of married couple households with children is enhanced.

Outside of the population living in private households are those living in institutions (hospitals, hostels, halls of residence, boarding schools, military establishments), who numbered about 820,000 in England and Wales in 1981 or about 1.7% of the total population.

The reconstruction phase of the microsimulation model proceeds by generating private household heads, the family head of constituent families, the spouse of the family head, their dependent children (if any), their non-dependent children (if any), and finally other adults. The reconstruction of private household members consists therefore of a series of nested loops, as set out in Figure 3. To these loops are added simpler loops for simulating the number of individuals in non-private households.

Once the members of the household have been generated, the size, composition and type of the household are then fixed and the household needs simply to be labelled appropriately. Each private household will occupy a housing unit, the characteristics of which are generated when all the household characteristics have been calculated, using conditional probablities dependent on household or household head attributes.

The number of private households and non-private household individuals to be simulated is fixed by the respective numbers recorded in the small area statistics of the 1981 Census. To update the population of the whole country would require the simulation of nearly 20 million private households and nearly 1 million individuals not living in private households. This is feasibly achieved only in small tranches, a small area at a time. There is no way that the interactions between more than 20,000 small area covering the whole country could be kept track of, so the model treats each small area as an individual open system with respect to migration.

2.2 The attributes

The UPDATE model aims to simulate most of the important population characteristics captured in the decennial census, and to improve on one or two of them, using national survey data. Table 1 sets out the attributes attached to individuals, Table 2 the attributes attached to households and Table 3 the attributes attached to housing units. Each list is discussed in turn.

The attributes in Table 1 are divided into four groups: the reference numbers which are generated automatically during the simulation, a set of demographic attributes, a set of socioeconomic attributes and a set of establishment attributes for individuals not in private households. Work to date has concentrated on the demographic set (see Rees, Clarke and Duley 1987 for a blueprint of how the socioeconomic and establishment attributes will be incorporated). Although most of the variables derive from census data, not all are available at small area level. For example, in arriving at the ethnic status groupings, national and district as well as small area data must be used. Note that the place of birth of the head of household is used to assign ethnic origin, and that an additional variable, I8, place of birth of the individual, is used to distinguish between first and second generation immigrants. Indicator I5, marital status, refers to legal status reported in the census by individuals whereas I6, marital condition, distinguishes living arrangements and attempts, using General Household Survey data, to incorporate informal unions of men and women, which have become much more important in developed country societies in the last two decades.

Table 2 sets out the list of household attributes that the model aims to incorporate. Reference number (HR1), type (H1) and size (H2) are generated in the reconstruction process loops laid out in Figure 3. The economic attributes in part are the sum of individual attributes but in part apply only to the households as a unit (H6, taxation and H7, net disposable income). Heavy use of non-census sources such as the UK's national Family Expenditure Survey will be made. The housing space attributes, H8 to H10, will be established using census data, and H11 from national survey sources. Attribute H12, tenure, comes direct from the small area census data.

In Table 3 the housing space attributes of Table 2 are replicated with one addition, an indicator, U1, of whether the housing unit is occupied. The list of housing units simulated will match that of households except for vacant or second home uits.

Vacant units will be available for occupancy by in-migrating households in years forward from the base year. The migration aspects of the model are discussed in section 5.2 of the paper.

2.3 The structured creation of household members

The reconstruction model is built on three foundations: firstly, the structured loops for head, spouse, dependent child, non-dependent child, other family adult, non-family adult and non-private household individual; secondly, the probability equations for assigning attributes to each of the different members; and thirdly, the techniques and data sources used to estimate the probabilities.

The general loop structure for household members has been illustrated in Figure 3. Figure 4 shows the stages involved in the generation of dependent children, which proved to require a considerable amount of fine tuning to avoid ending up with distributions of children by age that failed to match the age distributions provided directly in the small area statistics.

2.4 The probability equations

It will be useful here to show how the microsimulation model may be represented as a chain of probabilities. The probabilities apply in the generation of household heads and represent the target classifications of attributes. Some simplifications have been made for the purposes of presentation.

First, the head's age, sex and marital status are determined by sampling a joint probability array, P(a,s,m), where a stands for age, s for sex and m for marital status. The ethnic status of the head is then assigned using the conditional probability of ethnicity given age, sex and marital status,

$P(e|a,s,m)$. The place of birth of the head is computed from the conditional probability of birthplace given ethnic status, $P(b|e)$. To determine the family status of a household head is a two stage process. Firstly, the conditional probability of being a family member given ethnic status, age, sex and marital status, $P(f|e,a,s,m)$, is sampled, and secondly, those heads who are family members are tested for family head status by sampling the conditional probability of family headship given family membership, ethnic status, age and sex, $P(h|f,e,a,s,m)$. Then the marital condition of the household head is established on the basis of marital status, family type and status within the family by sampling $P(c|e,a,s,m)$.

The full chain of probabilities used to generate attributes I1 to I8 for the head of household is therefore

$P(c,h,f,b,e,a,s,m)$

$= P(a,s,m)\ P(e|a,s,m)\ P(b|e)\ P(f|e,a,s,m)\ P(h|f,e,a,s)$

$\times P(c|e,a,s,m)$

Similar probability chains are needed for other household members which use conditional probabilities dependent on one or more of the attributes assigned to the head of household.

The modelling problem is then one of estimating the probability arrays that enter the sampling chain, and the techniques for doing this are described in the next subsection of the paper.

2.5 <u>Probability estimation techniques</u>

To estimate the necessary probability arrays at the small area scale, use was made of several techniques.

2.5.1 Use of data from higher in the spatial hierarchy

The first technique involved disaggregating small area probability arrays through use of conditional probabilities computed from data tables available for units higher in the spatial hierarchy.

For heads, the reconstruction model begins with a joint probability array of age, sex and marital status available at the small area scale

$P^{s}(a4,m2,s1)$

where a stands for age, m for marital status and s for sex. The attached numbers indicate the different disaggregations of the attribute, with number 1 refering to the target set of categories while 2, 3 and 4 represent successively coarser disaggregations. At the district level a finer age disaggregation is available

$P^{d}(a3,m2,s1)$

from which it is possible to compute the conditional probability of a head being in an a3 category given membership of a4 age group, m2 marital status group and s1 sex group:

$P^{d}(a3|a4,m2,s1)$

so that the small are joint probability distribution is simply

$$P^{s}(a3,m2,s1) = \Sigma\, P^{d}(a3|a4,m2,s1)\, P^{s}(a4,m2,s1) \qquad (5)$$

Table 4 sets out the arithmetic involved. Sub-table (a) shows the initial joint probabilities (which sum to 1) for a Leeds postcode sector and sub-table (b) the more disaggregated (7 age groups rather than 4) joint probabilities for Leeds district. From these figures the conditional probabilities are computedd (sub-table (c)): note that no further deconsolidation is needed for two age groups. The conditional probabilities are applied using equation (5) and sub-table (d) shows the estimated joint probability distribution.

This technique is used on a myriad of occasions in the population reconstruction phase of the model to achieve the target categorization of the population set out in the attribute lists. In the example above, further disaggregation of age from classification a3 to classification a1 (single years of age) is needed via 5 year age groups (classification a2) and the marital status classification requires that the widowed/divorced category be split into its two components.

If anonymous individual census data or a public use sample were available, probability estimation would be much simplified. Special tabulations of the census could be requested but these are expensive, and involve long wait times.

2.5.2 Iterative proportional fitting

In some cases use of conditional probabilities derived from data for larger areas or the nation would result in estimates of a distribution across a particular dimension that violates what is known about that distribution at the small area scale.

For example, in the model we assign an age group to a wife dependent on the age group of her husband. The joint probability distribution of the ages of husbands and wives can be derived from national data (Table 5(a)). Note the clustering in the cells of and near to the principal diagonal, and the asymmetric nature of the other probabilities - some older men marry wives considerably younger than themselves, but older women do not marry very much younger men. For the small area, postcode sector LS6 1 in Leeds, the age distribution of married men and married women is known. Through iterative proportional fitting, the joint probability matrix (Table 5(a)) can be adjusted to fit the marginal constraints (Table 5(b) and 5(c)) to yield the estimated joint probability matrix. Stated formally, this procedure (after Fienburg, 1970; Clarke, 1986; Rees and Woods, 1986; Birkin, 1987) is as follows.

<u>Step 0</u>. Set the probabilities for the small area equal to those for the nation.

$$P^s{}_0(a_h, a_w) = P^n(a_h, a_w) \qquad (6)$$

where $P^s{}_0(a_h, a_w)$ is the initial estimate for the small area population of the joint probability of husbands' ages and wives' ages and $P^n(a_h, a_w)$ is the equivalent national probability.

<u>Step 1.</u> Constrain the estimates to the known marginal probabilities of husbands' ages:

$$P^s{}_1(a_h, a_w) = P^s{}_0(a_h, a_w) \times \left(P^s(a_h, *) \Big/ \sum_{a_w} P^s{}_0(a_h, a_w)\right) \qquad (7)$$

where $P^s(a_h, *)$ = known small area probabilities of husbands' ages.

<u>Step 2.</u> Constrain the estimates to the known marginal probabilities of wives' ages:

$$P^s{}_2(a_h, a_w) = P^s{}_1(a_h, a_w) \times \left(P^s(*, a_w) \Big/ \sum_{a_h} P^s{}_1(a_h, a_w)\right) \qquad (8)$$

where $P^s(*, a_w)$ are known small area probabilities of wives' ages.

<u>Step 3.</u> Test whether the estimates have converged:

$$\left| P^s{}_2(a_h, a_w) - P^s{}_0(a_h, a_w) \right| < e \qquad (9)$$

where e is a small probability, say .0001. If all probabilities satisfy inequality (9), the process is stopped. Otherwise, the initial probabilities are revised

$$P^{m}_{0}(a_h,a_w) = P^{m}_{2}(a_h,a_w) \qquad (10)$$

and the estimation returns to step 1. The iterations continue until the convergence criterion is satisfied. Sub-table (d) in Table 5 shos the converged estimates for postcode sector LS6 1 in Leeds.

The methods described here are used time and time again in different guises to estimate the probabilities needed in the population reconstruction. We now examine some results from the model and check on its reliability.

3. POPULATION RECONSTRUCTION: EXAMPLES AND TESTS

3.1 Simulated individuals

Table 6 lists the characteristics of 33 members of 11 households living in postcode sector LS6 1 in Leeds. This is an inner city neighbourhood to the north and northwest of the University of Leeds, characterized by a concentration of less well-off families, students and immigrants living in late nineteenth and early twentieth century terraced housing and in some urban renewal housing from the 1970s, predominantly of terrace and low rise flat design. There were some 3,875 households living in the area in 1981.

In the selection of households there are two single person households (numbers 11 and 3462), headed by a 19 year old single man and a 74 year old widowed lady respectively. Of the three two person households number 70 consists of a lone parent mother and her baby son, and number 3309 is made up of an elderly married couple. There are two three person households: in number 241 a single man, a divorced woman and her small daughter live together; number 887 consists of a young Irish married couple also with a daughter. Of the two four person households, number 2505 is the now atypical two parent-two child family. One child is non-dependent and was born to a very young mother (aged 17 at the time). The other four person household, number 32, is clearly a student "menage a quatre", sharing a rented house together. Such households are usually single sex in composition but not exclusively so. Finally, there are two larger households of five and six persons respectively, headed by men born in Pakistan and India. The Pakistani household is headed by the eldest male, living with his younger brother's family. The Indian household includes the head's widowed father as well as his own immediate family.

So, what do we make of these simulated households and their members? The compositions are certainly believable and one can imagine Barbara Taylor Bradford, the best selling authoress who has written about Leeds in earlier times, having a fine time reconstructing their personal histories. Later some formal tests of "believability" are carried out.

3.2 New crosstabulations

One of the advantages of the micro approach to multistate modelling is the ability to derive crosstabulations not available in the published statistics. Table 7 contains three such examples based on the mean of 5 simulation runs for postcode sector LS6 1. The first crossclassifies households by ethnic status of head against number of persons in the household; the second crossclassifies ethnic status of head against household composition (a reduced version of the full classification); the third tabulates persons by age group, sex and marital condition. This last table makes an estimate of the numbers living under the new informal arrangement of cohabitation. This represents new information about the population, though, of course, it should be treated with caution because critical interdepencies may have been omitted from the reconstruction phase of the model.

3.3 Test of the simulations

To test the robustness of the population reconstruction phase of the UPDATE model, crosstabulations were computed from the mean of 5 simulations that exactly matched statistics available in the small area tabulations but which had not been directly used in the model. Table 8 sets out Z-scores for the population classified by sex, marital status (two categories) and age (5 year age groups). The Z-scores were computed as

$$Z_i = (OP_i - PP_i) / ((1 - PP_i)/n)^{0.5} \qquad (11)$$

where OP_i is the observed and PP_i the preddicted probability for category i and the divisor expression is an estimate of the standard error. If Z-scores fall outside the values + or - 1.96, we cannot be confident at the 95% level that the observed and predicted probabilities are drawn from the same population. For the 99% confidence level the critical Z-value is 2.80.

In virtually all cases (93 out of 106) the differences between predicted and observed probabilities are within the 95% confidence band. All of the significant differences are in the unmarried category and concentrate in the 20-24, 25-29 and 75+ age groups for men in both areas, and in the 20-24, 25-29 and 70-74 age groups for women in the inner city zone, and in the 10-14 age group in the suburban zone. Each of these differences can be traced to the lack of local conditional probabilities at key points in the reconstruction model. The more local probabilities that can be used in the model the better. The overall fit of the model predictions is very good - the indices of dissimilarity between observed and predicted numbers over the population categories in Table 8 are 4.5 and 2.8 for the inner city and suburban zones respectively on a scale of 0 (perfect fit) to 100 (no similarity). The corresponding correlation coefficients are 0.986 and 0.993. In the model blueprint (Rees, Clarke and Duley, 1987) the suggestion was made that model outputs could be constrained to fit such observed distributions as that used in the Table 8 comparison exactly. However, in practice, this proved to be computationally unworkable and a small model bias must be accepted.

The next section of the paper describes how the simulated individuals are exposed to the risks of mortality and fertility.

4. UPDATING: VITAL COUNTS

4.1 The locational system

So far the exact nature of the small area units being used in the model has not been specified. These are, in fact, postcode sectors or aggregations of the unit postcodes which the British Post Office uses to mechanize the sorting of mail. Each unit postcode covers a small number of addresses (or properties) and amalgamation from the unit level, e.g. LS16 7HR, to the sector level, e.g. LS16 7, produces areas containing circa 3,000 households and 7,500 individuals. The census is conducted using a different geography:

enumerators are assigned enumeration districts (EDs) of circa 500 households and 1,400 persons. The census geography has been matched to the postal geography by Pin Point Ltd by assigning enumeration districts to postcode sectors based on proximity of their respective areal centroids, each ED being assigned to the nearest postcode sector.

However, the statistics on vital events are available only down to census ward level, wards being aggregations of 30 to 60 EDs. The procedure adopted has been to estimate the relevant vital rates for wards, assume that these rates apply to each of the constituent EDs, and then to apply a weighting function based on the ED populations when aggregating EDs from different wards to form postcode sector equivalent populations. The postcode sector vital rate is thus a weighted average of the rates for the wards with which it intersects areally.

4.2 The estimation of mortality probabilities

The microsimulation model uses mortality rates for each single year of age (period-cohort) transition and each sex. The rate is updated for each year. The target variable to be estimated is:

$m^{ws}_{aa+1}(y)$ = probability that a person in ward w in age group a at the start of year y dies in that year before attaining age a+1.

To estimate these mortality probabilities we have available national deaths by single years of age, ward deaths by eleven aggregate age groups (roughly ten year age groups) and the corresponding populations at risk for the base year. National population estimates are produced for each mid-year, as are district population estimates which are used to crudely factor the

1981 ward population. So the variables available for the estimation are:

$D^{ns}_a(y)$ = deaths in year y in nation n by sex s and single year of age a

$D^{ws}_c(y)$ = deaths in year y in ward w by sex s and coarse age group c

$P^{ns}_a(y)$ = mid-year population in year y for nation n by sex s and age a

$P^{ws}_b(y)$ = population in year y in ward w by sex a and age group b.

The nation used is, in this application, England and Wales. The age classification is as follows: a = single years of age from 0 to 74, and 75 and over; b = single years of age 0-24, 5 year age groups 25-29, ... ,70-74 and 75 and over; and c = 0,1-4,5-14, ... , 65-74, 75 and over. Classification a is the target one for the UPDATE model, classification b is the best that can be achieved for ward population estimates, and classification c is the one used to report the ward deaths.

The estimation then proceeds as follows.

National death rates are computed for age disaggregation a

$$d^{ns}_a(y) = D^{ns}_a(y) / P^{ns}_a(y) \qquad (12)$$

and age disaggregation b

$$d^{ns}_b(y) = \sum_{a \in b} D^{ns}_a(y) \Big/ \sum_{a \in b} P^{ns}_a(y) \qquad (13).$$

The ward level death rates are then the national death rates adjusted to satisfy ward deaths figures:

$$d^{ws}_a(y) = d^{ns}_a(y) \times \left[D^{ws}_c(y) \Big/ \sum_{b \in c} d^{ns}_b(y)\, P^{ws}_b(y) \right] \qquad (14)$$

The mortality probabilities are then computd, shifting the age-time plan from period-age (used so far) to period-cohort (required in any demographic projection model):

$$m^{we}_{aa+1}(y) = 1 - [(1 - 0.5\ d^{we}_{a+1}(y))/(1 + 0.5\ d^{we}_{a}(y))\ ]\ (15).$$

Slightly more complicated estimation procedures applying the same technique are required for the first two ages and the last two ages, and these are described in Duley (1988, Chapter 5).

The resulting mortality probabilities are obviously less smooth for the small area than for the nation, but do reflect actual events that have occurred there. For projection purposes, the procedure is aggregated over 5 years to produce less variable, more reliable long run probabilities, which are modified using national trends.

4.3 <u>The estimation of maternity probabilities</u>

In the model women are exposed to the risk of giving birth in each year expressed as a maternity probability. New infants are introduced into the family and household by sampling from a probability distribution of the numbers of births per maternity. The probabilities are concentrated, of course, at one birth, but the model allows for twins, triplets, quadruplets and quintuplets to be born using a sharply declining probability function. Births are sexed using local probabilities of a male or female infant.

Maternity probabilities are dependent, like mortality, on sex and age, but also, rather critically, on marital status and ethnic origin. National information on the age and marital status classification of births, on the age and ethnic group classification of births is combined with district level data on births by age, marital status and by ethnic group, together with total births by ward. The available fertility and population at risk statistics are defined in Table 9 together with the target variable $pm^w(a,m,e)$. The following procedure is used to estimate the target variable for use in the microsimulation model.

Step 1: computation of national birth rates

Iterative proportional fitting is used to estimate the array $B^n(b,m,e)$ from marginals $B^n(b,m)$, which is an aggregation of $B^n(a,m)$, together with $B^n(b,e)$ and $B^n(e,m)$ (see Duley, 1988, Chapter 5 for the details and section 2.5.2 for the general method). The conditional probability that a birth is to a mother of age n, given she is in age group b (computed from the $B^n(a,m)$ array), is then applied to yield an array of births classified by single years of age, marital status and ethnic group

$$B^n(a,m,e) = \sum_b p^n(a|b,m,e)\ B^n(b,m,e) \qquad (16).$$

Equivalent methods are used to estimate the female population at risk, and national birth rates are computed

$$b^n(a,m,e) = B^n(a,m,e)\ /\ F^n(a,m,e) \qquad (17).$$

Step 2: computation of district birth rates

An initial estimate of the number of births by age group, marital status and ethnic group was computed using conditional probablities derived from the national estimation

$$B^d(b,m,e) = P^n(e|b,m)\ b^d(b,m) \qquad (18).$$

This initial estimate was then adjusted to known district marginals, $B^d(b,m)$ and $B^d(e)$ using IPF techniques again. District populations at risk, $F^d(b,m,e)$ were estimated from marginal arrays $F^d(b,m)$, $F^d(b,e)$ and $F^d(e,m)$ using IPF. The district birth rates could then be estimated as

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

for a e b (19)

Step 3: computation of ward birth rates

IPF was again used to estimate the female population at risk, $F^w(b,m,e)$ from marginal arrys $F^w(b,m)$ and $F^w(e)$, so that the following estimate could be made of ward birth rates

$$b^w(a,m,e) = b^d(a,m,e) \left[B^w / \sum_b \sum_m \sum_e b^d(b,m,e) F^w(b,m,e) \right]$$

for a e b (20)

Step 4: conversion of ward birth rates to maternity probabilities

The national ratio of maternities by mother's age to births by mother's age was used to adjust ward birth rates

$$m^w(a,m,e) = b^w(a,m,e) \left[M^n(a) / B^n(a) \right] \qquad (21)$$

and these period-age rates were converted to period-cohort probabilities thus

$$pm^w(a,m,e) = 0.5 \, (1 + s^w(a)) \, m^w(a,m,e) \qquad (22).$$

In the simulation, if women experience a maternity after this probability distribution has been sampled, then the number-of-births per maternity distribution is sampled to produce one or more new infants. The infants are sexed by sampling the local probabilities of a male or female birth. When the ward level maternity probabilities are updated for years after the base year, rather cruder estimates of the ward populations at risk have to be made, but the estimates will reflect both changing national and changing local fertility behaviour.

4.4 Illustrative results

Do these estimation procedures produce sensible results at the ward level? Figure 5 graphs birth rates for Headingley ward in Leeds in 1981 for married and unmarried women by age for the eleven ethnic groups employed. The ordering of the age groups in terms of level of fertility accords with what is known about the national picture (Population Statistics Division, OPCS, 1988). The estimation procedures for mortality and maternity probabilities can be regarded as either a method for adjusting national rates to local perturbations, or as a method for smoothing the erratic fluctuations at the small area scale using national schedules. These are classic strategies in multistate population analysis.

5. UPDATING: FURTHER PROCESSES

Sections 2 through 4 of the paper report on model development that has already been accomplished. This section sketches out the intentions for incorporating marital transition, migration and socioeconomic transition into the UPDATE model, intentions which are part way to fulfilment.

5.1 Marriage and cohabitation; divorce and dehabitation

Clarke (1986) developed procedures to handle marriage and divorce in his microsimulation model of the population of Yorkshire and Humberside. The present model uses some of the same techniques, particularly those of pair matching through examining the ages of potential partners (the probability matrix for which was used as illustration in section 2 of the paper). However, UPDATE differs from previous work in two respects: firstly, the phenomenon of cohabitation (a man and a woman living together as husband and wife without having married in a civil or religious ceremony) is incorporated, and secondly, the marriage market defined is an open one. Individuals are free to migrate to find a spouse or cohabitee, and are free to attract a spouse or cohabitee to the area, resulting in the addition of an in-migrant to the simulated population.

The algorithm for marriage and cohabitation and its consequent events is set out in Figure 6. There are three stages in the algorithm: (i) testing each individual in the population for eligibility (this essentially excludes those under 16, and those who are members of married couples); (ii) testing the eligible individuals for marriage or cohabitation; and (iii) matching selected individuals with partners and working out the consequent location decisions for both partners. Note that one route through the algorithm allows cohabiting couples to marry - no matching or migration is involved.

Outlining the matching process, the gender of the base partner is first randomly selected. Selection is random to avoid sex bias in the matching process. Based on the 5 year age group a_1 of a base partner, an imaginary partner (spouse or cohabitee) is generated of age group a_2. Then a match of this imaginary person with a locally available person from the relevant partner pool is attempted. If a local partner fitting the age-sex requirement is found a match is made and the two individuals are united and removed from their respective partner pools. If there is no suitable partner locally available, then an interface with the wider marriage/cohabitation market outside the small area is necessary. This interface results in either the export of the local individual (and their dependent and non-dependent children where present) and their removal from the pool, or the import of an exotic partner of the required age and sex. The decision to import or export individuals to satisfy the partner search is based on the annually adjusted local in- and out- migration propensities of individuals by age and sex. Gross figures are used so that there will be a combination of in- and out-migration motivated by marital transitions.

This open marriage/cohabitation model not only matches the real situation more closely than would a locally closed process, but also has the additional attraction of allowing a total allocation principle to be included whereby all prospective partners are successfully matched either internally or externally without resort to pool balancing and overcomes the problem of matching of the last few couples in a closed system in which considerable age gaps may be forced to appear.

When the match has been completed the individuals' marital status and condition attributes are updated. The newly formed couple must then be allocated to a household. The procedure adopted depends on the involvement of migration in the matching process. The out-migration of the local partner removes them from the resident population so by-passing the allocation problem, although the departure requires that the origin household's attributes be updated and this may or may not add a vacant housing unit to local stocks. In the case of in-migration of an exotic partner, the new partnership is formed initially in the local partner's existing household, the attributes of which are updated. When both partners are locally resident the match will result in movement. Either or both the partners can move, however, the modelling implications of such an event are complex. To combat this problem, the algorithm can be split up into two stages.

In the first stage, within the marriage/cohabitation model, movement is restricted to one of the two partners so that the couple or family is formed in an existing household, with no new household being established. The determination of who moves uses a rule-based approach involving a combination of tenure, density measures (persons per room) and household composition preference favouring smaller single unit households. If the two households are the same in every respect the female's household is adopted.

In the second stage, these new couples/families are flagged for migration as a wholly moving unit to be tested in the migration model in the existing simulation period. Couples made up of a local and an exotic partner are similarly flagged. This reflects the often unsuitable and temporary nature of the initial housing situation of newly formed couples. Such flagged households are then given a higher than average probability of moving as a couple or family. This step is supported by the finding of Grundy and Fox (1984), who discovered very high migration rates among women in their first year of marriage, rates that rapidly fall off to normal levels with increasing duration of marriage. Although the data set dates back to 1970-71, the trend is so clear and the phenomenon so ongoing that a suitable adjustment factor for 1981 and subsequent migration probabilities should be computed.

The modelling of pair dissolution is more straightforward than that of marriage, in that no matching algorithm is required since it is this match that is split through the divorce or dehabitation process. However, the implications of pair dissolution are equally complex. The probability of divorce can be computed from the Marriage and Divorce Statistics (Series FM2 OPCS) based on the combination of age groups of the two marriage partners, eligibility being based on the presence of married couple(s) within a household, as distinct from simply having marital status married, for in such cases a person may be separated. This combined age indicator is required due to the variation in the age distribution of divorcees by sex. In the absence of data on the termination of cohabiting relationships, divorce rates are adopted as an estimate of the rates of dehabitation. The joint effect of lack of available data and the complex set of possible consequences for household dynamics means that certain assumptions concerning the outcome of the divorce transition must be made, such as, (i) the former husband sets up a new single person household, (ii) the former wife and any child(ren) remain in their present dwelling unit with updated attributes reflecting their new circumstances, and (iii) both former partners now become eligible for remarriage.

5.2 Migration

Crucial to the success of the UPDATE simulation is the ability to represent migration correctly in the model, and to make reasonable estimates of the rates or probabilities involved. The small area whose population is being simulated is conceived of as an open system. People living there are free to leave and new people may be attracted in, but both flows are dependent on conditions and events within the area, a multiarea model having been rejected as infeasible at an early stage of design.

Three processes are involved in this view of migration (Rees, Clarke and Duley, 1987, p.22): (i) the mobility process or the propensity of residents in an area to move; (ii) the destination selection process or the propensity of residents to select a destination within or without the area; and (iii) the in-migration process, the propensity of vacant housing units and housing units vacated by mobile residents in the area to attract households or individuals from outside the area. The mobility process can be regarded as determined in part by the demographic attributes of the population (age, sex, marital status, the role of the latter having been discussed earlier in this section) and in part by socioeconomic and housing attributes, of which tenure is particularly important. The destination selection and in-migration processes are determined by the number of housing vacancies in the area and the size of the area in relation to typical migration fields.

This discussion suggests the following strategy for incorporating migration into the microsimulation model.

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

(ii) 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.

(iii) 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; other will out-migrate. The successful matching of moving household with housing units in the housing pool will in part depend on a probabilistic 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 to produce some within area migrants and some out-migrants.

(iv) 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.

(v) 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 or common reference numbers.

Is there sufficient information about migration to implement these steps? The 1981 Census provides information on the number of persons and households moving into each enumeration district. National tables from the 1981 Census provide migration data aggregated by age and sex of head of household, and detailed data are available at national and district scale of the age and sex attributes of migrants. This information will be used to establish mobility probabilities for the base year (cf. section 4). The mobility probabilities will be updated using information on current migration produced from the National Health Service Central register (NHSCR) for Family Practitioner Committee (FPC) areas (districts within former metropolitan counties, shire counties and groupings of London boroughs). Application of a time series index is essential because of the rise that has occurred in migration activity since 1981 (Stillwell, Boden and Rees, 1988). It would be valuable to also include a local area adjustment factor for a current year, based on a source such as the electoral register, but there are difficulties in identifying in-migrants to and out-migrants from an electoral area with current registers (although this was possible in the 1970s).

Having established the necessary mobility propensities for wholly moving households and individuals, it will be necessary to allocate the migrants generated to housing units within and without the area. Inter-ward migration data are now available from the 1981 Census (via the MATPAC information sysytem) which can be used to calculate, for a particualr area, the distance function of migration. Combined with knowledge of the small area's size this should make possible a division of migrating residents into out-movers and within area movers. The within area movers would compete for housing units in the small area, but movers out of the area would release units for occupation by in-migrants.

In-migrating households move to such vacated housing spaces, to newly constructed dwellings, to dwellings vacant at the start of the time interval, and in-migrating individuals will move to spaces within existing households. Knowledge of the number of new dwellings constructed will thus be needed. This is usually available at local government level, but could also be estimated, perhaps, from electoral register data.

The comments above have been tentative, but sufficient to indicate that development of a migration algorithm in the UPDATE microsimulation model will be feasible with existing data sources in the UK.

5.3 Socioeconomic transitions

To date little progress has been made in including individual socieconomic or household economic attributes in the model as attention has been focused on the demographic components. There is no difficulty in deriving probability distributions at the small area scale from the 1981 Census for these indicators, nor in using national surveys (General Household Survey) or information systems (such as the National Online Manpower Information System) to establish updated cross-sectional characteristics. What will be more difficult will be to establish the transition probabilities between economic states. Some suggestions were put forward for their estimation in Rees, Clarke and Duley, 1987. We hope to report further in a future paper.

6. CONCLUSIONS

This paper has been an interim report on an ambitious project. The project has, as its aim, the simulation of the dynamic processes that affect residential populations living in small areas. Sufficient progress has been made and sufficient experience gained with the techniques of data estimation to look forward with confidence to the completion of the project. An eclectic approach to data sources has been forced upon us and innovative approaches to data estimation have had to be developed. By adopting the microsimulation approach we have had to think hard about the real events that people's lives and real determinants of people's behaviour, which is the only way, ultimately, to make progress in understanding socio-spatial systems.

REFERENCES

Betson, D., Greenburg, D. and Kasten, R. (1980) A microsimulation model for analysing alternative welfare reform programs for better jobs and income. In <u>Microsimulation models for public policy analysis, Vol.1., distributional impacts</u>. Academic Press, New York.

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. (1988) SYNTHESIS: a synthetic spatial information system for urban and regional analysis with methods and examples. Environment and Planning A, forthcoming.

Clarke, M. (1986) Demographic processes and household dynamics: a microsimulation approach. In Woods, R.I. and Rees, P.H. (eds.) Population structures and models. George Allen and Unwin, London.

Clarke, M. and Holm, E. (1987) Microsimulation methods in spatial analysis and planning. Geografiska Annaler, 69B, 2, 145-164.

Clarke, M. and Spowage, M.E. (1985) Integrated models for public policy analysis: an example of the practical use of simulation models in health care planning. Papers of the Regional Science Association, 55, 25-45.

Clarke, M. and Wilson, A.G. (1986) A framework for dynamic comprehensive urban models: the integration of accounting and microsimulation approaches. Sistemi Urbani, 2/3, 145-177.

Corner, I.E. (1987) Developing the DOE household projection method. Building Research Establishment, Department of the Environment, Note N.83/87. Watford, Herts.

Duley, C. (1988) A model for updating households in small areas in intercensal periods. Ph.D. thesis in preparation, School of Geography, University of Leeds.

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

Galler, H. and Wagner, G. (1986) The microsimulation model of the Sfb 3 for the analysis of economic and social policies. In Orcut, Merz and Quinke (eds.) Information research and resource reports, Volume 7. North Holland.

Giesbrecht, F.G. and Fiew, L. (1969) Demographic microsimulation model POPSIM II. Manual for programs to generate vital events. Technical report No.5, Project SV-285, Research Trianlgle Institute.

Hagerstrand, T. (1952) The propagation of innovation waves. Lund Studies in Geography, B, 4, Lund.

Hagerstrand, T. (1967) On Monte Carlo simulation of diffusion. In Garrison, W. and Marble, D.F. (eds.) Quantitative geography Northwestern University Press, Evanston, Illinois.

Holm, E. (1984) To locate education, employment and residence. Gerum B, Department of Geography, University of Lund.

Holm, E., Makila, K. and Oberg, S. (1985) Time geographic concepts and individual mobility behaviour. Paper presented at the Philadelphia Regional Science Association meeting, November.

Holm, S. and Oberg, S. (1984) Migration in micro and macro perspective. In Scandinavian Population Studies, 6, 1. Stockholm.

Kain, S. and Apgar, W. (1985) Housing and neighbourhood dynamics: a simulation study. Harvard University Press, Cambridge, Mass.

Martin Voorhees Associates and John Bates Services (1981) Developing the migration component of the official subnational population projections. Final Report prepared for the DPRP3 Division, Department of the Environment, Martin Voorhees Associates.

OPCS (1984) Census 1981. Household and family composition tables. HMSO, London.

OPCS (1988) Mid 1987 population estimates for England and Wales. OPCS Monitor, Series PP1, 88/1. OPCS, London.

Orcutt, G., Caldwell, S. and Wertheimer II, R. (1976) Policy exploration through microanalytic simulation. The Urban Institute, Washington.

Population Statistics Division, OPCS (1988) Live births in 1987. Population Trends, 53, 35-40.

Rees, P.H., Clarke, M. and Duley, C. (1987) A model for updating individual and household populations. Working Paper 486, School of Geography, University of Leeds.

Rees, P.H. and Willekens, F. (1986) How the Dutch and the English adopted multiregional models for subnational population projection. Working Paper 472, School of Geography, University of Leeds. Forthcoming in Stillwell, J.C.H. and Schoelten, H. (eds.) The comparative geography of the UK and the Netherlands. Reidel, Dordrecht.

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

Rogers, A. (1976) Shrinking large-scale population projections models by aggregation and decomposition. Environment and Planning A, 8, 515-541.

Rogers, A. and Willekens, F. (eds.) (1986) Migration and settlement: amultiregional comparative study. Reidel, Dordrecht.

Stillwell, J.C.H., Boden, P. and Rees, P.H. (1988) Internal migration change in the UK: trends based on NHSCR movement data, 1975-6 to 1985-6. Working Paper 510, School of Geography, University of Leeds.

Table 1 The list of individual attributes.

Label	Description	Categories or description
Reference numbers		
IR1	household number	sequence number of household generated
IR2	family number	family number within household
IR3	individual number	number of person in family
Demographic attributes		
I1	status within household	head, non-head, not member of private household
I2	status within	family head, spouse, dependent child, non-dependent child, non-family adult
I3	age	single years to 75, 75 and over
I4	sex	male,female
I5	marital status	single, married, widowed, divorced
I6	marital condition	independent, member of married couple, member of cohabiting couple
I7	ethnic status	origin: British, Irish Rep., AuCaNz, Pakistani, Indian, Bangladeshi, West Indian, E. African, African Rem., NC Remainder, Other EC, Other European, Rest of world
I8	place of birth	inside UK, outside UK
Socioeconomic attributes		
I9	economic activity	active, retired, student, inactive
I10	occupation	socioeconomic groups I to XVII with V split into V.1 and V.2
I11	employment status	full-time, part-time, self-employed, unemployed
I12	industry	agriculture, forestry and fishing; energy and water; manufacturing; construction; distribution and catering; transport and communication; other services
I13	earned income	income by virtue of work
Establishment attributes		
I14	establishment status	staff, residents, not in establishment
I15	establishment type	hotel or hostel; children's home; old people's home; psychiatric hospitals; other hospitals; schools & colleges; prison establishments; lodging houses; other establishments

Table 2 The list of household attributes.

Label	Description	Categories or description
Reference number		
HR1	household number	household sequence number
Composition and size		
H1	type based on composition	see Figure 2 for the detailed classification of types A to Q
H2	number of persons in household	1,2,3,4,5,6,7 or more
Economic attributes		
H3	earned income	sum of individual earned incomes
H4	benefits	sum of benefits received by members
H5	unearned income	income from savings and investments
H6	taxation	taxes, the result of applying tax rules to H3, H4 and H5
H7	net disposable income	income less taxation
Housing space attributes		
H8	tenure	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	1,2,3,4,5,6,7 or more
H10	amenities	exclusive use of all amenities; lacking inside w.c.; lacking hot water; lacking inside w.c. and hot water
H11	value or rent	either value of housing unit or rental paid for the unit
Household durables		
H12	number of cars	no car, one car, two or more cars

Table 3 The list of housing unit attributes.

Label	Description	Categories or description
Reference number		
UR1	housing unit number	sequence number of housing unit
Housing space attributes		
U1	occupance	occupied first home, occupied second home, vacant
U2	tenure	as for attribute H8 (Table 2)
U3	rooms	as for attribute H9 (Table 2)
U4	amenities	as for attribute H10 (Table 2)
U5	value or rent	as for attribute H11 (Table 2)

Table 4 An illustration of the procedure to extend attribute categories

(a) The joint probability of age, sex and marital status of head of household in Leeds postcode sector LS6 1 in 1981. (1)

age group		males			females	
(A4)	married	single	wid/div	married	single	wid/div
16-29	0.06344	0.17090	0.00466	0.01476	0.11471	0.00777
30-44	0.09296	0.04298	0.01476	0.01346	0.01994	0.01528
45-64/59	0.10953	0.02382	0.01787	0.00647	0.01036	0.02304
60/65 +	0.06758	0.01217	0.02279	0.00621	0.03599	0.08855

(b) The joint probability of age, sex and marital status of head of household in Leeds District in 1981. (2)

age group		males			females	
(A3)	married	single	wid/div	married	single	wid/div
16-19	0.00087	0.00164	0.00001	0.00028	0.00206	0.00002
20-24	0.02218	0.01129	0.00035	0.00284	0.00874	0.00105
25-29	0.05043	0.00990	0.00200	0.00400	0.00585	0.00407
30-44	0.19534	0.01246	0.01035	0.01064	0.00719	0.01914
45-64/59	0.24052	0.01452	0.01773	0.00709	0.00749	0.02946
65/60-74	0.08102	0.00528	0.01350	0.00487	0.01393	0.07106
75 +	0.02743	0.00229	0.01255	0.00156	0.01080	0.05617

(c) The conditional probability of being in age group (A3) <u>given</u> age group (A4).

age group			males			females	
(A4)	(A3)	married	single	wid/div	married	single	wid/div
16-29	16-19	0.01184	0.07184	0.00424	0.03933	0.12372	0.00389
	20-24	0.30185	0.49452	0.14830	0.39887	0.52492	0.20428
	25-29	0.68631	0.43364	0.84746	0.56180	0.35136	0.79183
30-44	30-44	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
45-64/59	45-64/59	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
65/60 +	65/60-74	0.74707	0.69749	0.51823	0.75739	0.56328	0.55852
	75 +	0.25293	0.30251	0.48177	0.24261	0.43672	0.44148

(d) The 'new' joint probability of age, sex and marital status in postcode sector LS6 1 for an extended age grouping (A3) adopting Leeds district information.

age group		males			females	
(A3)	married	single	wid/div	married	single	wid/div
16-19	0.00075	0.01228	0.00002	0.00058	0.01419	0.00003
20-24	0.01915	0.08451	0.00069	0.00589	0.06021	0.00159
25-29	0.04354	0.07411	0.00395	0.00830	0.04031	0.00615
30-44	0.09296	0.04298	0.01476	0.01346	0.01994	0.01528
45-64/59	0.10953	0.02382	0.01787	0.00647	0.01036	0.02304
65/60-74	0.05049	0.00849	0.01181	0.00470	0.02027	0.04946
75 +	0.01709	0.00368	0.01098	0.00151	0.01572	0.03909

Source: (1) computed from SAS Table 26
(2) computed from County Report Table 35 p.58

Table 5 The estimation of the age combinations of married couples.

(a) National data : age of husband by age of wife. [1]

Age of husband	Age of wife 16-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75+
16-19	.0010	.0004	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
20-24	.0043	.0262	.0040	.0005	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
25-29	.0010	.0278	.0466	.0072	.0010	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000
30-34	.0003	.0062	.0392	.0627	.0088	.0014	.0004	.0001	.0000	.0000	.0000	.0000	.0000
35-39	.0001	.0013	.0074	.0431	.0481	.0061	.0011	.0003	.0001	.0000	.0000	.0000	.0000
40-44	.0000	.0004	.0020	.0092	.0381	.0416	.0063	.0011	.0003	.0001	.0000	.0000	.0000
45-49	.0000	.0002	.0007	.0025	.0091	.0358	.0388	.0066	.0012	.0003	.0001	.0000	.0000
50-54	.0000	.0001	.0002	.0009	.0027	.0095	.0345	.0397	.0078	.0015	.0004	.0001	.0001
55-59	.0000	.0000	.0001	.0004	.0009	.0025	.0091	.0341	.0410	.0089	.0016	.0004	.0001
60-64	.0000	.0000	.0001	.0002	.0003	.0008	.0023	.0098	.0315	.0314	.0068	.0012	.0003
65-69	.0000	.0000	.0000	.0001	.0001	.0003	.0008	.0026	.0092	.0285	.0287	.0063	.0011
70-74	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008	.0025	.0066	.0196	.0209	.0047
75+	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008	.0018	.0051	.0137	.0242

(c) Local data : age of married men [2]

.0028
.0873
.1374
.1177
.0709
.0871
.0680
.0857
.0869
.0713
.0769
.0530
.0551

.0161	.1443	.1464	.1060	.0602	.0831	.0729	.0879	.0717	.0600	.0670	.0440	.0404

(b) Local data : age of married women [2]

(d) Local estimate of age of husband by age of wife. [3]

Age of husband	Age of wife 16-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75+
16-19	.0019	.0008	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
20-24	.0111	.0681	.0073	.0006	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
25-29	.0020	.0586	.0699	.0063	.0005	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000
30-34	.0005	.0117	.0523	.0486	.0037	.0007	.0002	.0001	.0000	.0000	.0000	.0000	.0000
35-39	.0002	.0025	.0099	.0337	.0205	.0034	.0005	.0001	.0000	.0000	.0000	.0000	.0000
40-44	.0002	.0014	.0043	.0117	.0265	.0372	.0048	.0009	.0002	.0001	.0000	.0000	.0000
45-49	.0001	.0004	.0012	.0028	.0054	.0275	.0253	.0044	.0006	.0002	.0001	.0000	.0000
50-54	.0001	.0003	.0006	.0013	.0021	.0096	.0298	.0350	.0052	.0010	.0003	.0001	.0001
55-59	.0000	.0002	.0003	.0006	.0008	.0029	.0088	.0336	.0309	.0067	.0017	.0004	.0001
60-64	.0000	.0001	.0002	.0003	.0003	.0009	.0023	.0099	.0244	.0241	.0071	.0012	.0004
65-69	.0000	.0001	.0001	.0001	.0001	.0004	.0009	.0030	.0080	.0219	.0333	.0074	.0016
70-74	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0007	.0018	.0048	.0191	.0204	.0057
75+	.0000	.0001	.0001	.0001	.0000	.0001	.0001	.0003	.0006	.0014	.0054	.0145	.0324

Source:
1 - computed from Household & Family Composition Table 24 (OPCS (1984))
2 - computed from SAS Table 21 for postcode sector LS6 1
3 - computed from national dataset (a) constrained to local marginals (b) and (c)

Table 6. Thirty-three Typical Tykes

Sequence number	Reference numbers h'hold	fam.	ind.	Status in h'hold I1	family I2	Age I3	Sex I4	Marital status I5	condition I6	Ethnic status I7	place of birth I8
:	:	:	:	:	:	:	:	:	:	:	:
29	11	0	1	head	other	19	M	single	I	UK	UK
:	:	:	:	:	:	:	:	:	:	:	:
69	32	0	1	head	other	17	M	single	I	UK	UK
70	32	0	2	non-head	other	22	M	single	I	UK	UK
71	32	0	3	non-head	other	17	M	single	I	UK	UK
72	32	0	4	non-head	other	21	F	single	I	UK	UK
:	:	:	:	:	:	:	:	:	:	:	:
108	70	1	1	head	head	20	F	single	I	W. Indies	UK
109	70	1	3	non-head	child	1	M	single	I	W. Indies	UK
:	:	:	:	:	:	:	:	:	:	:	:
267	144	1	1	head	other	26	M	single	I	Pakist'i	abroad
268	144	1	2	non-head	head	22	M	married	MC	Pakist'i	abroad
269	144	1	3	non-head	spouse	26	F	married	MC	Pakist'i	abroad
270	144	1	4	non-head	child	0	F	single	I	Pakist'i	UK
271	144	1	5	non-head	child	1	F	single	I	Pakist'i	UK
:	:	:	:	:	:	:	:	:	:	:	:
401	241	1	1	head	other	21	M	single	CC	UK	UK
402	241	1	2	non-head	head	28	F	divorcd	CC	UK	UK
403	241	1	3	non-head	child	1	F	single	I	UK	UK
:	:	:	:	:	:	:	:	:	:	:	:
1401	887	1	1	head	head	28	M	married	MC	Irish	abroad
1402	887	1	2	non-head	spouse	26	F	married	MC	Irish	UK
1403	887	1	3	non-head	child	2	F	single	I	Irish	UK
:	:	:	:	:	:	:	:	:	:	:	:
1734	948	0	1	head	other	26	F	married	CC	UK	UK
1735	948	0	2	non-head	other	22	M	single	CC	UK	UK
:	:	:	:	:	:	:	:	:	:	:	:
2073	1540	1	1	head	head	34	M	married	MC	Indian	abroad
2074	1540	1	2	non-head	spouse	33	F	married	MC	Indian	abroad
2075	1540	1	3	non-head	child	3	F	single	I	Indian	UK
2076	1540	1	4	non-head	child	2	F	single	I	Indian	UK
2077	1540	1	5	non-head	child	8	M	single	I	Indian	UK
2078	1540	1	6	non-head	other	72	M	widowed	I	Indian	abroad
:	:	:	:	:	:	:	:	:	:	:	:
6819	2505	1	1	head	head	54	M	married	MC	UK	UK
6820	2505	1	2	non-head	spouse	43	F	married	MC	UK	UK
6821	2505	1	3	non-head	child	8	F	single	I	UK	UK
6822	2505	1	4	non-head	non-dep	24	M	single	I	UK	UK
:	:	:	:	:	:	:	:	:	:	:	:
7891	3309	1	1	head	head	71	M	married	MC	UK	UK
7892	3309	1	2	non-head	spouse	65	F	married	MC	UK	UK
:	:	:	:	:	:	:	:	:	:	:	:
7967	3462	0	1	head	other	74	F	widowed	I	UK	UK
:	:	:	:	:	:	:	:	:	:	:	:

Source: extracted from a simulation run for postcode sector LS6 1

Table 7. New crosstabulations: examples for postcode sector LS6 1

(a) Ethnic status of head by number of persons in the household.

Ethnic status of head	number of persons in household										
	1	2	3	4	5	6	7	8	9	10+	Total
UK	1140	902	425	247	132	49	12	4	1	0	2912
Irish	62	43	17	9	8	3	2	0	0	0	145
Au/Can/NZ	4	3	1	1	0	0	0	0	0	0	10
Pakistani	45	47	29	17	18	11	6	4	3	3	182
Indian	82	56	30	20	14	9	6	2	1	1	221
Bangladeshi	2	4	1	0	1	0	0	0	0	0	7
W. Indian	10	19	11	6	4	3	2	0	0	0	56
E. African	8	10	8	4	3	2	1	0	1	0	37
Rest Africa	14	10	5	2	1	0	0	0	0	0	31
Rest NC.	16	24	15	12	8	4	1	0	0	0	80
Other E.C.	23	13	6	5	2	0	0	0	0	0	51
Other Euro.	11	7	4	3	1	1	0	0	0	0	28
Rest World	49	36	15	6	4	2	1	0	0	0	114
All heads	1465	1175	567	333	197	84	31	13	6	4	3875

(b) Ethnic status of head by household composition.

Ethnic status of head	household composition									
	no family		1 family							2+ families
			married couple				lone parent			
	1 person	2+ persons	with no children	dept children only	dept and non-dept	non-dept only	dept children only	dept and non-dept	non-dept only	
UK	1140	517	461	326	80	111	177	16	68	10
Irish	62	28	16	11	7	5	9	1	5	0
Au/Can/NZ	4	2	2	0	0	0	0	0	0	0
Pakistani	45	50	15	38	9	2	12	0	1	10
Indian	82	47	23	26	7	5	16	1	5	11
Bangladeshi	2	4	0	0	0	0	1	0	0	0
W. Indian	10	6	7	11	3	2	10	2	3	0
E. African	8	7	7	9	1	1	2	0	0	4
Rest Africa	14	10	3	1	0	0	2	0	0	0
Rest NC.	16	13	17	22	4	2	3	0	0	1
Other E.C	23	13	4	4	2	1	2	0	1	0
Other Euro.	11	6	3	4	1	1	0	0	0	1
Rest World	49	29	16	8	3	3	4	1	1	1
All heads	1465	734	578	461	116	135	237	21	86	38

(c) Age by marital condition by sex of persons aged 15+.

Age group	males			females		
	independent	married couple	cohabiting couple	independent	married couple	cohabiting couple
15-19	291	4	3	298	18	5
20-24	680	98	23	513	191	32
25-29	493	210	24	282	224	18
30-34	152	173	7	89	155	6
35-39	90	100	6	58	77	3
40-44	71	118	3	53	118	3
45-49	56	95	3	55	99	3
50+	411	578	13	738	494	12
total	2244	1376	82	2086	1376	82

source: computed from the mean of 5 simulation runs.

Table 8 A comparison of observed and predicted populations in two Leeds postcode sectors: Z scores.

age group	LS6 1 an inner city zone		LS19 7 a suburban zone	
	Married		Married	
	Males	Females	Males	Females
15-19	0.00	-0.59	0.00	-0.28
20-24	-1.40	-1.50	1.50	0.70
25-29	1.83	0.69	-0.70	0.00
30-34	0.00	0.32	1.12	0.45
35-39	-0.10	0.64	0.78	1.12
40-44	-0.09	-0.27	0.59	0.92
45-49	0.20	0.39	0.30	-0.37
50-54	0.64	0.09	0.36	0.43
55-59	0.71	-0.10	0.12	0.36
60-64	0.49	-0.64	0.20	0.00
65-69	0.86	0.20	0.35	-0.24
70-74	0.68	-0.73	0.41	0.09
75+	-1.10	-0.14	-0.44	-0.58
	Unmarried		Unmarried	
	Males	Females	Males	Females
0-4	2.28*	1.69	1.46	-2.20*
5-9	0.44	0.00	-0.39	0.69
10-14	-1.70	-1.56	4.79**	4.46**
15-19	0.18	-0.06	-1.69	-1.23
20-24	-5.02**	-2.95**	-2.27*	-0.82
25-29	3.14**	3.04**	-3.34**	-1.45
30-34	-1.05	-0.22	-1.55	-0.39
35-39	-0.94	0.41	-1.90	-0.28
40-44	-1.18	-0.44	-1.65	-0.35
45-49	0.75	1.04	0.19	0.49
50-54	0.60	1.12	1.08	0.00
55-59	0.98	1.09	0.56	0.50
60-64	0.14	1.34	0.42	0.83
65-69	0.29	1.83	0.15	0.85
70-74	0.70	2.12*	0.00	0.53
75+	3.26**	1.66	2.09*	0.78

Notes:
1. z-score = (observed probability - predicted probability)/standard error
The observed number is derived from Small Area Statistics Table 21
2. * = z-score beyond 95% confidence limits (but within 99%)
3. ** = z-score beyond 99% confidence limits

Table 9 Fertility and population statistics used in the estimation of ward maternity probabilities.

Variable	Definition
Fertility statistics	
$B^n(a,m)$	births in nation n by age a and marital status m
$B^n(b,e)$	births in nation n by age group b and ethnic status e
$B^n(e,m)$	births in nation n by ethnic group e and marital status m
$M^n(a)$	maternities in nation n by age a
$B^d(b,m)$	births in district d by age group b and marital status m
$B^d(e)$	births in district d by ethnic group e
B^w	total births in ward w
Female populations at risk (PAR)	
$F^n(a,m)$	PAR in nation n by age a and marital status m
$F^n(b,e)$	PAR in nation n by age group b and ethnic status e
$F^n(e,m)$	PAR in nation n by ethnic status e and marital status m
$F^d(b,m)$	PAR in district d by age group b and marital status m
$F^d(b,e)$	PAR in district d by age group b and ethnic status e
$F^d(e,m)$	PAR in district d by ethnic status e and marital status m
$F^w(b,m)$	PAR in ward w by age group b and marital status m
$F^w(e)$	PAR in ward w by ethnic status e
Also used	
$s^w(a)$	survival probability for ward w by age a for women (estimated previously, see section 4.2)
Subscript classifications	
a	single years of age 15 to 49 inclusive
b	15, 16-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49
m	married, unmarried
e	UK, Irish, Old Commonwealth, African NC, Caribbean NC, Bangladeshi, Indian, NC Remainder, Pakistani, Other EC, Rest of world (NC = New Commonwealth)
Target variable	
$pm^w(a,m,e)$	maternity probability for ward w for mothers of age a, marital status m and ethnic status e

Notes:

1. The subscripts in brackets all apply to the female population only.

Model phase	Input
POPULATION RECONSTRUCTION	National, regional, local Census data, national surveys
UPDATING TO THE PRESENT	National, regional, and local demographic events, National surveys, Local administration data
PROJECTION INTO THE FUTURE	Trended events, new scenarios, alternative policies

Figure 1. The general structure of UPDATE.

Statistics for England & Wales for 1981
(Source: OPCS,1984,Table 1)

	Number of families.	Couple status	Child status	Others status	Type	% of households	% of persons	Average size
All private households	No family		1 person		A	21.7	8.1	1.0
			2 or more persons		B	4.8	3.9	2.2
	One family	Lone parent	Non-dependent child(ren) only	With other(s)	C	0.5	0.6	3.3
				Without other(s)	D	3.1	2.6	2.2
			Dependent and non-dependent child(ren)	With other(s)	E	0.2	0.4	5.0
				Without other(s)	F	0.8	1.2	3.8
			Dependent child(ren) only	With other(s)	G	1.2	1.7	3.9
				Without other(s)	H	2.4	2.5	2.8
		Married couples	No children	With other(s)	I	1.5	1.7	3.1
				Without other(s)	J	24.4	18.1	2.0
			non-dependent child(ren) only	With other(s)	K	0.6	0.9	4.4
				Without other(s)	L	7.5	9.4	3.4
			Dependent and non-dependent child(ren)	With other(s)	M	0.4	0.9	5.7
				Without other(s)	N	5.5	10.1	4.9
			Dependent child(ren) only	With other(s)	O	1.1	2.1	5.0
				Without other(s)	P	23.3	34.1	4.0
	Two or more families				Q	0.9	1.9	5.6
				Totals :		100.0	100.0	2.8
				Number :		17.71	48.81	

Figure 2 The household classification

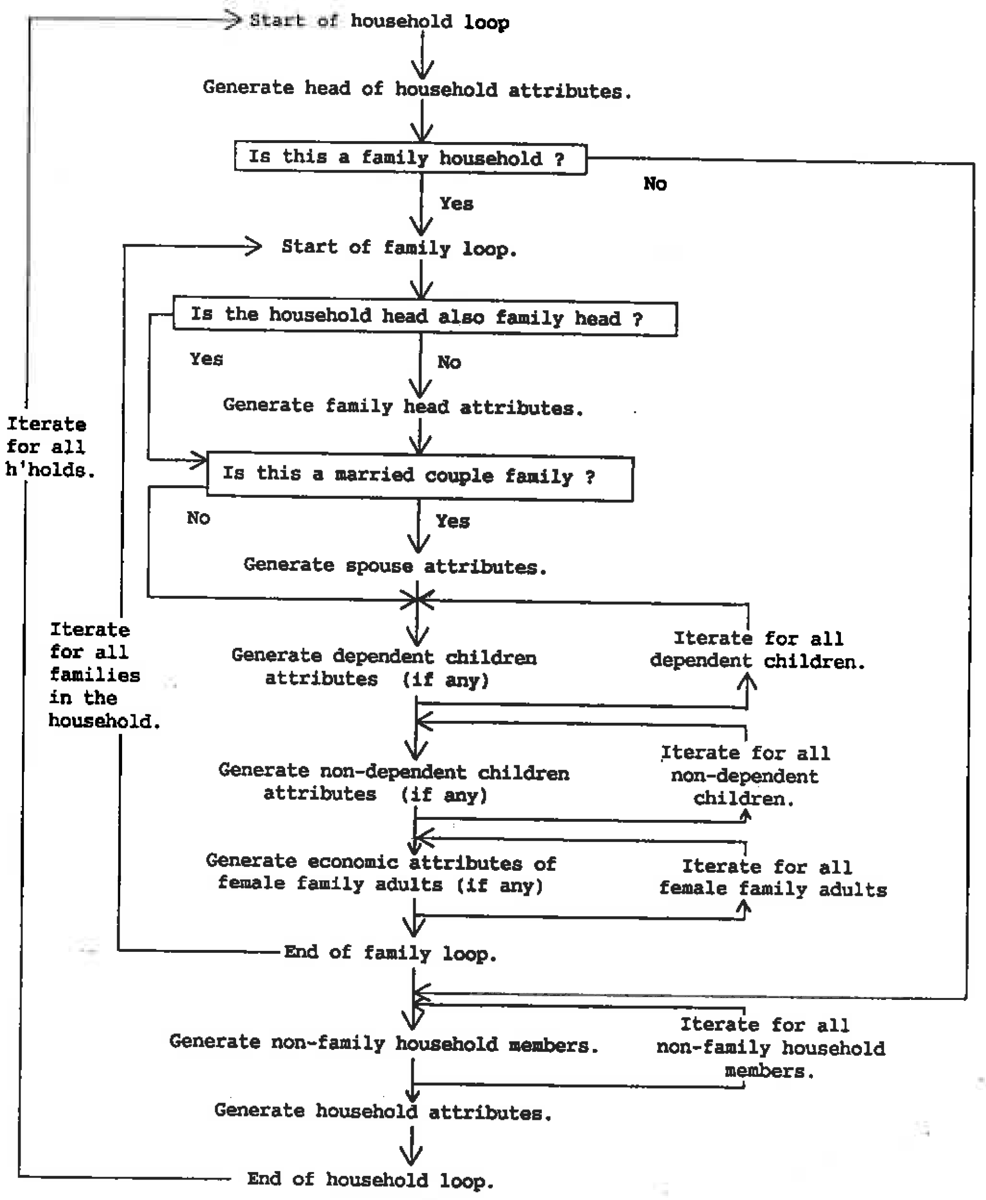

Figure 3 Household, family and individual loops in the reconstruction process.

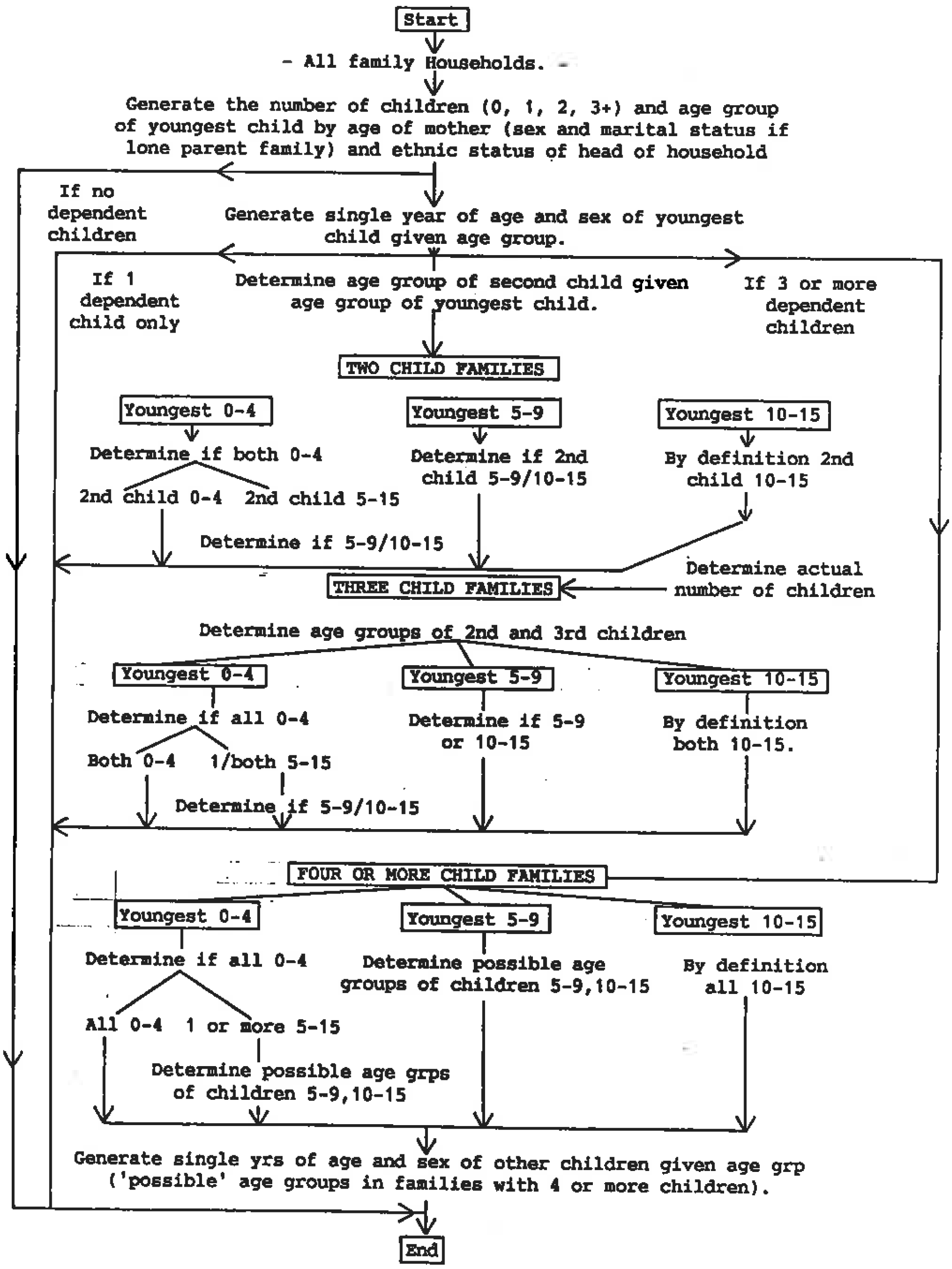

Figure 4 Stages in the generation of dependent children.

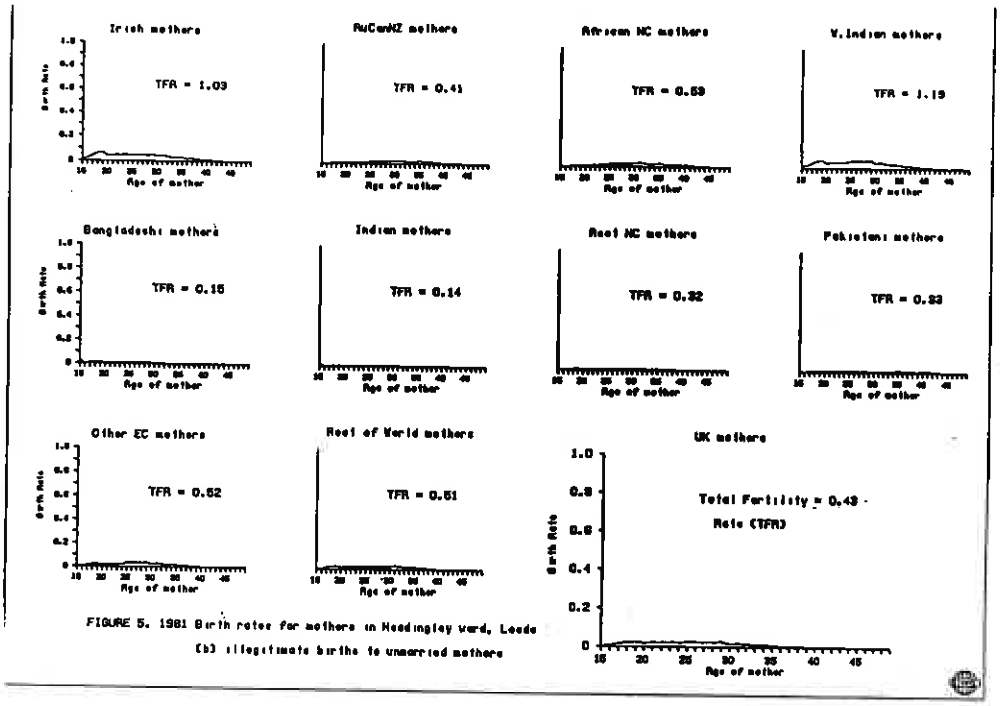

FIGURE 5. 1981 Birth rates for mothers in Headingley ward, Leeds
(b) illegitimate births to unmarried mothers

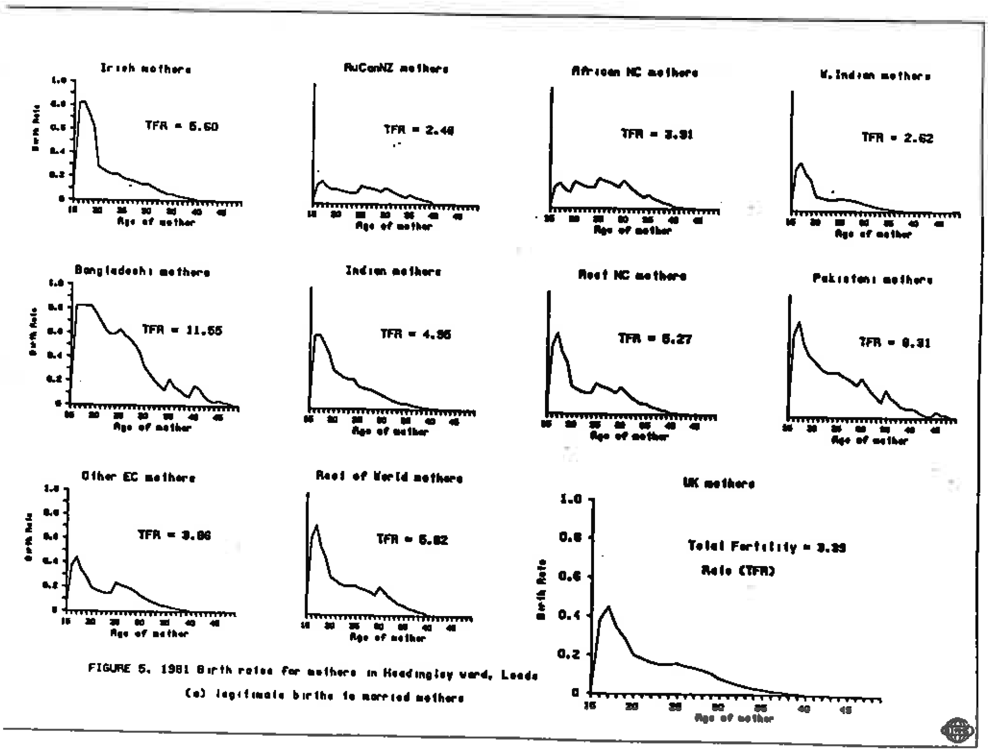

FIGURE 5. 1981 Birth rates for mothers in Headingley ward, Leeds
(a) legitimate births to married mothers

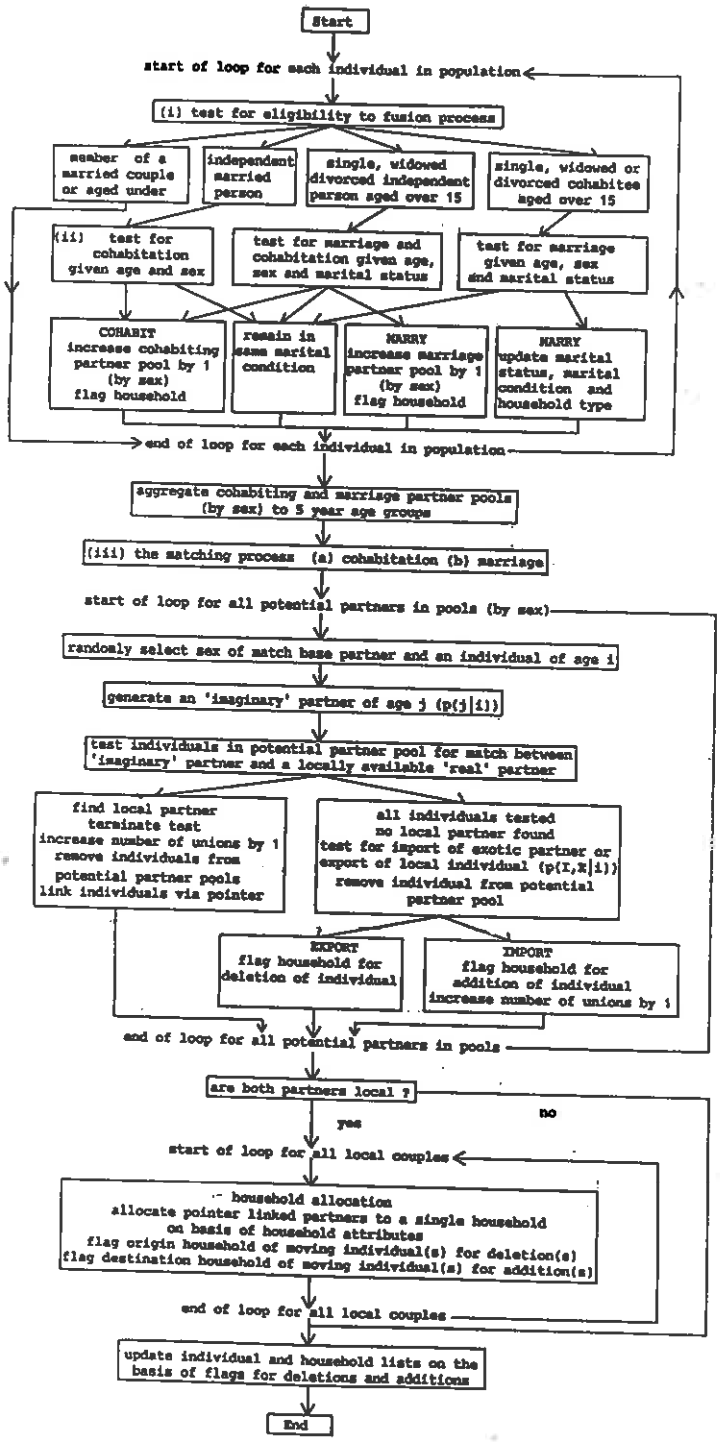

Figure 6. Stages in the simulation of cohabitation and marriage.