Working Paper No. 241

DYNAMICS OF FAMILY SIZE AND COMPOSITION : A COMPUTER SIMULATION STUDY WITH REFERENCE TO RURAL INDIA

N. Krishnaji
Chandan Mukherjee

[This work was carried out as part of the Research Programme under the International Training Programme in Population and Development, funded by the UNFPA, New York.]

Centre for Development Studies
Thiruvananthapuram
December 1990

1. Introduction

1.1 The focus of this paper is on household size variations in rural India in the post-Independence period. Attention will however be paid to several related issues such as the age-sex composition of households and, more importantly, the structural complexities - described by the variation from simple or nuclear to extended and joint households - that have a direct bearing on size.

1.2 In this paper we set out the context in brief (section 2), explain the simulation model designed to answer the questions posed (section 3), and present some results based on a few experimental runs of the model (section 4). Although these results have a readily seen substantive content, they are discussed here mainly to indicate the scope of the model in future work.

2. The Context

2.1 Data on household size variations in rural India are mostly based on sample surveys, with the population Censuses providing some additional information. These variations are usually presented by classificatory variables such as the household ownership or operational holding, and per capita consumer expenditure. However, the reference is always to households rather than families, households being defined by the extended residence criterion, conforming thus broadly to units of production and consumption. The distinction between households and families - howsoever defined - are well known and indeed of obvious relevance to the study of variations in size. However, in rural India, it is only the large landowners who have attached farm servants residing in, and therefore being counted in surveys as members of, households. More importantly, these 'unrelated' residents constitute only a small fraction of household membership.

2.2 It can be seen from Table 1 that for the population as a whole in 1971 only about 19 per cent of the households `report' having attached farm servants, with a mean of 1.86 servants per reporting household.

Table 1: Average Household Size by Landholding-All India Rural

Size Class of Holding (in Acres)	1971-72				1982
	Average Household Size according to		%Household reporting attached servants in operational holdings	Average No. of servants per reporting holding	Average Household Size according to Ownership Holding
	Ownership holding	Operational holding			
0.00	3.73	4.15	-	-	4.14
0.00 - 0.05	4.72	4.67	6.81	1.30	4.79
0.05 - 1.00	4.92	4.83	11.15	1.41	5.16
1.00 - 1.25	4.88	4.90	10.71	1.46	5.18
1.25 - 2.50	5.35	5.26	16.17	1.51	5.50
2.50 - 5.00	5.66	5.78	19.63	1.68	6.01
5.00 - 7.50	6.27	6.36	20.84	1.83	6.57
7.50 - 10.00	6.58	6.79	25.13	2.19	7.00
10.00 - 12.50	6.75	6.88	27.63	2.19	6.87
12.50 - 15.00	7.18	7.18	31.86	2.23	7.37
15.00 - 20.00	7.59	7.51	33.85	2.19	7.46
20.00 - 25.00	7.77	7.86	38.14	2.36	8.08
25.00 - 30.00	7.60	7.79	41.59	2.32	8.06
30.00 - 50.00	8.01	8.08	55.76	2.72	8.48
50.00 +	9.07	9.23	62.36	3.42	8.88
All	5.34	5.34	18.96	1.86	5.43

Sources: National Sample Survey, India, Report No.215 and Sarvekshana, India, October 1987.

This means that for the population as a whole the average number of servants per household is 0.35, whereas the mean household size is 5.34, i.e. attached servants constitute about 6.5 per cent of the population. In the largest landholding class, 62 per cent of the households report servants, with a mean of 3.42 servants per reporting household and, therefore, a mean of 2.13 per household in this landholding class. The average size of household in this class is 9.23. Thus the large differences in average household size between small and big farmers, ranging from 4 to 9 remain even after `unrelated' members are not counted: the range will still be from 4 to 7. Accordingly, the concepts household and family can be used interchangeably without severely invalidating conclusions derived from empirical data analysis.

2.3 In rural India, mean household size is found to systematically increase with the size of landholdings, and of assets or wealth, more generally. A large mean size is the characteristic of large landowning families; the agricultural labour and small farmer families have a much smaller mean size, although a considerable variation in family size exists within each landholding class. It is with explanations for this contrast between, say, a mean size of 4 or less in the small landholding category and a mean in excess of 8 in the very large landholding class that we are mainly concerned (see Table 1).

2.4 It must be mentioned here that the observed positive correlation between family size and wealth appears to be universal in the sense that it holds across both space - for the different rural subregions of India - and time - as readily verified by the different landholding surveys conducted during the post-Independence period. It bears further generalisation to the extent that in several contemporary third world countries as well as some in pre-industrial Europe family size and wealth are found to be correlated likewise (Krishnaji, 1980)[1]. Because of the apparent universality of this correlation in countries predominantly agrarian in economic character, it will be tempting to infer that the explanation must also be universal. The search for an universal explanation is bound to fail, however, because family size variations are induced by a number of demographic parameters - sometimes interrelated in complex ways -and more than one possible configuration of demographic parameters may produce the same type of correlation. Thus, given that undivided (joint or extended) families are generally large in size simply because they contain more than one simple family unit, it is difficult to interpret the differences in the demographic structures - for example in fertility and mortality rates - underlying the observed differences in mean size.

2.5 The difficulty is compounded by the absence of survey estimates of demographic rates such as of fertility, mortality and nuptiality separately for classes of agricultural labour, small farmer and big farmer households. The published data in sample surveys classified by size of landholding generally refer to household size and its age-sex composition. In addition, the composition of the household membership according to 'relationship to the head' of the family have been tabulated in one population census. The latter data sets, relating to the 1961 census, containing much detail on household size and composition indeed show that household structure increases in complexity with the size of landholding. Households with large landholdings thus contain a far higher proportion of married relations (of the head) - indicating a higher incidence of joint families of some kind or another - than do the families with small landholdings.

One such data set is presented in Table 2. These data relate to Kerala, one of the major states of Indian Union. Data for other states reflect a similar pattern. It can be seen that married sons (of heads) constitute 2.2 per cent and other married relations (presumably, brothers of heads, because married women live with their husbands' families) 6.7 per cent of the household population in the smallest landholding class of less than one acre. These percentages systematically increase as the landholding size increases: in the largest size class they are 4.2 and 10.7 per cent respectively. These data suggest, therefore, that at least one factor underlying the observed family size-wealth positive correlation is the tendency for families with substantial wealth to remain undivided for a longer time than for poor families. Thus the rate at which new households are set up from old ones - we call this the <u>splitting or partitioning rate</u> in this study - also contributes to family size variations, apart from birth and death rates. Accordingly, one of the objectives of the study is to examine the sensitivity of family size to variations in parametric configurations consisting of fertility and mortality schedules, rates and rules of marriage and partitioning of families.

Table 2: Composition of Sample Households by Relationship to Head of Family Classified by Size of Land Cultivated: Kerala, 1961.

		Composition of Population (Per Cent)					
Size of Holding (acres)	Average Size of Household	Heads of Household and Spouses	Married Relations: Sons	Married Relations: Others	Never Married, Widowed, Divorced/ Separated Persons	Unrelated Persons	Total
Less than 1.0	5.70	30.7	2.2	6.7	59.9	0.3	100.0
1.0 - 2.4	6.42	27.9	3.1	7.9	50.2	0.9	100.0
2.5 - 4.9	7.16	25.3	3.7	9.2	59.6	2.2	100.0
5.0 - 7.4	7.48	24.0	4.1	9.9	58.4	3.6	100.0
7.5 - 9.9	8.00	23.9	3.8	10.9	58.2	4.9	100.0
10.0 or more	8.13	21.9	4.2	10.3	57.5	6.1	100.0

<u>*Source*</u>*: Social and Cultural Tables, Census of India, 1961.*

2.6 We must refer in this context to the studies initiated by Laslett (1972)[2] on the size and structure of households in pre-industrial Europe. These showed, among other things, that in several rural communities the mean size varied within a very narrow range (typically between 4 and 6). A computer simulation study done in that context has established that generally the mean size is more sensitive to rules of household formation than to demographic variation embodied in fertility and mortality (Burch, 1972)[3]. While our simulation model extends the results of earlier work, it is designed with somewhat different aims.

Firstly, we are interested in the demographic picture in rural India roughly depicting the period 1950-80, different from that of pre-industrial Europe. Secondly, the studies of Laslett et al. (Wachter, 1978)[4] focus on a particular hypothesis, viz the stem family hypothesis, for explaining variations in the household structure rather than in household size whereas our concern here is more with family size, its distribution and the extent to which it is influenced by rates of household formation captured in a fairly general manner without reference to family structures in an explicit manner.

3. The Simulation Model

3.1 Overview of the Model: The simulation model computes the annual transition of a given population with characteristics such as age, sex and marital status, according to a specified set of demographic parameters. This makes it possible to take a given population through a pre-specified demographic regime over a certain period of time and record its consequences at the end of the period. Further, by repeating the exercise on the same initial population with different demographic experiences it is possible to examine how different demographic configurations influence the outcomes.

A population is initially specified. This population corresponds, in our experiments, to the beginning of the time point of the period over which the dynamics is to be studied. The population is specified by a list of households, and a list of individuals within each household with the following characteristics -

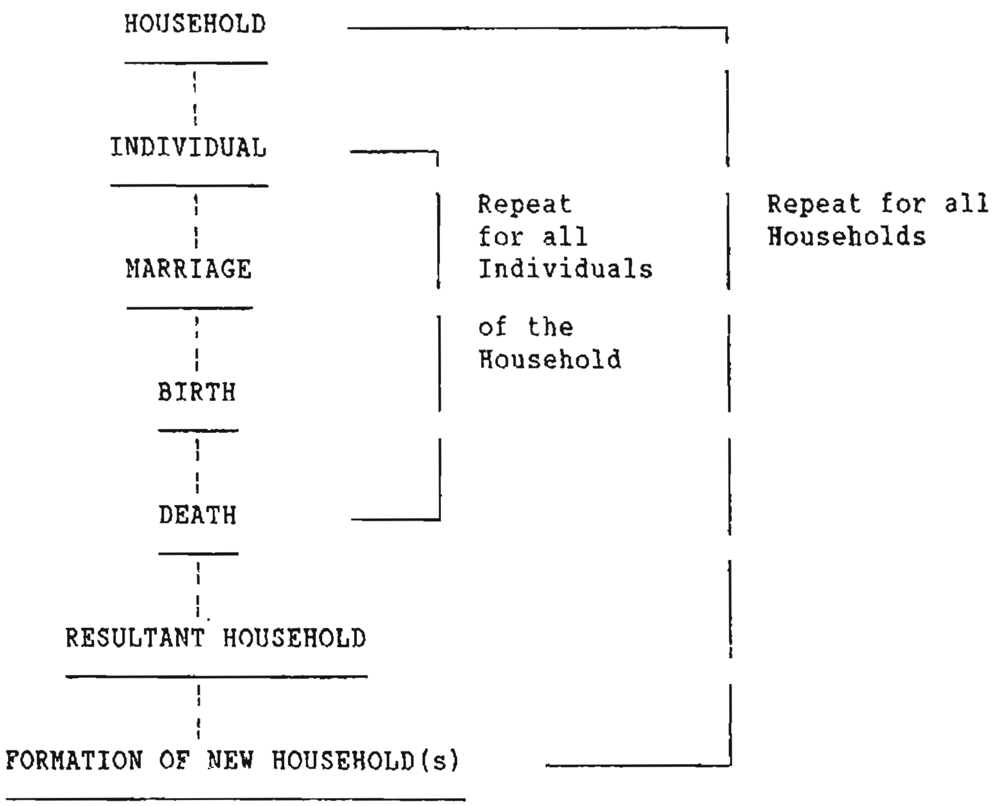
1. Identity of the household
2. Identity of the individual
3. Generation of the individual
4. Age of the individual
5. Sex of the individual
6. Marital Status of the individual
7. Identity of the parent/guardian
8. Identity of the marital partner
9. Year when the last child was born (for ever-married females)
10. Number of children ever born (in case the individual is an ever-married female)

The identities listed in items 1,2,7 and 8 are specified by serial numbers in the relevant list. The generation number is meaningful only in the relative sense i.e. children will have this number set to one plus that of the parents. The identity number of the parent or guardian (seventh characteristic listed above) needs some explanation. It is zero for the eldest member of the household. It is the identity number of the husband in the case of wife. When a child is born, this code for the child is the same as that of the mother of the child. Subsequently, this code changes due to deaths in the household over time. We shall refer to this number as the 'dependency code'. For example, if the mother dies, the dependency code will change to the identity number of the father, if the latter is alive. More on this when we discuss the death module.

The unit of time is a year. Every year demographic events (like marriage, birth, death) occur to each individual according to pre-specified probabilities; new households are formed out of the initial (at the beginning of the year) set of households according to a given household formation rule and probability. These events are recorded and the new list of population and households corresponding to the beginning of the next year is constructed.

The model actually works out the annual transition household by household. Each individual in a household is subjected to the probabilities of demographic events in the following sequence - marriage, birth, death. (Note that an event may be inappropriate for a given individual. For example, birth of a child is impossible to a male, but this only means that the probability will be set to zero in this case. By subjecting an individual to the probability of an event we mean that the event will occur to the individual with the given probability. For example, if the probability of an unmarried female at age 16 to get married is 0.032 then an arbitrary unmarried female at age 16 will get married in the current year with the same probability).

After subjecting each individual to the probabilities of different vital events, the resultant changes in the household are then recorded, or in other words, a new list of the individuals is prepared (with the particulars as stated before) excluding those who die during the year and females who marry and move out, and including those who move into the household after marriage to a male within the household. The resultant household is now subjected to the probability of breaking up (partitioning) forming new household(s) according to certain pre-specified rules. Schematically, the steps involved in computing the annual transition can be presented as follows.



3.2 Perhaps the best known and a very elaborate microsimulation model of this kind is SOCSIM developed by Hammel et al.(1976)[5]. In comparison, our model is simple in its construction, has minimal computer requirements and easy to implement for applications of a certain type. The program was written in BASIC language and implemented on an IBM PC/AT compatible machine. We have christened the program FAMSIM and henceforth it will be referred to by this name. The major simplifications in FAMSIM as compared to SOCSIM will be pointed out while we describe the model in detail in the Appendix.

What follows is a brief description of the different sections (modules) of the model. More detailed descriptions of each section (module) of the model are presented in the Appendix. The simulation model specification is completely empirical. The probabilities are specified, as described in the appendix, either directly or through mathematical functions producing reasonable distributions of events, obtained by trial and error.

3.3 Marriage Module: Marriage probabilities are applied to males of age between 15 and 49 and to females of age between 10 to 44 years (inclusive of the limits). The difference between the ages of the husband and the wife has been kept constant at 5 years. The simulation model is an open one so far as marriage is concerned. When a female marries, she drops out of the household membership list, whereas when a male marries, the wife is added to the list. This conforms to actual practice in India, with rare exceptions.
No divorce or separation is allowed but re-marriages are allowed.

3.4 Birth Module: Child birth probability depends on the woman's age, marital status and parity i.e. total number of children she has already borne. There is no explicit specification for birth spacing except for a very simple assumption that a married woman cannot bear more than one child in two successive years.

3.5 Death Module: Age-specific death probabilities are based on Life Tables. When a death occurs the dependents, if any, of the deceased will have to be assigned a new guardian. The following conventions are specified in the model for this purpose.

a. If the deceased is a married male there is no change required since dependents are linked to the wife of the married male.

b. If the deceased is a married female then all her dependents will be transferred to the husband.

c. If the deceased is not currently married (i.e. unmarried or widow or widower) then all the dependents will be

   (i) transferred to the wife of the eldest member of the household if the eldest member is a married male,
   or (ii) transferred to the eldest member if the eldest member is not currently married but is aged above 15 years,
   or (iii) dropped out of the population if there be no one aged above 15 years left in the household i.e. we assume that the children will join some other household.

Notice that a given individual in the household may be identified with several individuals of the household through the dependency code i.e. several individuals may have the same dependency code which identifies this particular individual. All these 'dependents' are not necessarily this particular individual's own children. The dependency code has been used to keep track of relationship between individuals within the household in a simplified way, which is essential for determining the membership of new households formed out of the parent household over time.

3.6 Formati n of New Households: This section of the model is based on a specification in which the probability of splitting or partitioning of a household is a function of the size of the parent household, apart from a convention explained below. The probability is set to zero for households with less than two living couples. The probability increases as the size of the household increases reaching unity at or before size 20. In other words, a household is not allowed to grow beyond size 20. This limit is arbitrary but adequate for our purpose.

Let us define 'Critical Size' as the size of a household when it breaks up or partitions. Given the probability specification (of partitioning) one can compute the mean critical size. It is the average size of the households at the time of their partitioning. We shall, in the subsequent discussion, use the mean critical size corresponding to a given specification of partitioning probability as a measure of the proneness of households to partition.

When a household breaks up, it has to be partitioned (forming new households) according to a certain given rule.

The eldest member of the household has been designated as the head of the household. Each new household forms with a couple, their children and dependents. Thus, the household formation rule specifies which are the couples in the parent household who form as many new households. Further details on the rule are presented in Appendix.

3.7 Generating the Initial Population: The first problem is to specify the initial population embodying characteristics of the past demographic regimes. It is not possible to do this by listing out a population off hand. We take the help of the model itself. We proceed as follows: Start with a population of what we might call a certain number of 'Adam-Eve' couples. We specified 200 couples - each pair of age 20 (male) and 15 (female) years, each without any children. In other words, we specify a list of 200 households each consisting of a husband and a wife only of age 20 and 15 years respectively. We then specify a demographic regime which corresponds to the fertility, mortality and nuptiality patterns of India during the decade of 1930s and take this population with this same set of demographic parameters over 150 years under an arbitrarily specified probability for new household formation. The choice of the population of 200 couples is not of any consequence in the subsequent analysis because the evolution of this population under constant vital rates becomes rapidly independent of the initial composition. The terminal population (at the end of 150 years) is taken as one which should roughly match with the population of India in 1941. Notice that we started with an arbitrary probability specification for formation of new households. This specification affects only the distribution of size and composition of the resultant households and not the structure of population which depends on the vital rates only. By trying out the above exercise with a few specifications we could arrive at one that produces an average household size of 4.73, reasonably close to that of rural India in 1941. We must stress at this point that the attempt to correspond the demographic rates and other characteristics of the population with that of specific periods in India is only for ensuring that we work with a reasonable range of the parameters and characteristics of population; <u>it is not crucial to the analysis we set forth with the simulation results</u>; <u>viz, that of the sensitivity of mean size to variations in demographic parameters within a plausible range</u>.

Table 3: Specification for Generation of the Initial Population and Some of the Statistics Obtained in the Process

INPUT SPECIFICATION	Simulation Years	
	1-150	151-160
1. Nuptiality	1931-41 India	1941-51 India
2. Mortality	Coale-Demeney Life Table with e0(m) = 33 e0(f) = 32	Coale-Demeney Life Table with e0(m) = 33 e0(f) = 32
3. Marital Fertility (TMFR)	7.5	7.5
4. Mean Critical Household Size	10.0	10.0
OUTPUT	Years 141-150 (1931-40)*	Years 151-160 (1941-50)*
1. Average Crude Birth Rate	45	43
2. Average Crude Death Rate	33	31
3. Average Household Size	4.73	4.83

* Roughly corresponding to the period.

The computed population is further evolved over another 10 years with the specification of vital rates corresponding to the 1941-51 period in India to produce what should be our initial population for simulation experiments.

This population should roughly correspond to that of India in 1951.
The input specifications for the model and some of the rates obtained in the simulation exercise while producing the initial population are given in Table 3.

3.8 Subsequent Evolution: We study how this initial population evolves under different demographic regimes over a period of thirty years. The choice of a period of thirty years of evolution needs clarification. The common practice is to derive the stable distribution corresponding to a given set of demographic parameters to analyze the results. (We have followed this procedure to generate the initial population). The derivation of stable distribution is feasible when the given parameters remain constant over time. The period over which the convergence to the stable distribution takes place is not of interest for substantive analysis of the results derived. Stable distribution theory is merely an analytical tool. What is of substantive interest in the present context is the impact of different demographic parameters and household formation rules on the household size and its distribution over a certain period. For this purpose it is enough to choose a period over which the impact of the specified factors (demographic etc.) is discernible. It is not possible to analytically derive such an impact over a specified finite period. A computer micro-simulation model has precisely this advantage. We have chosen 30 years (or three decades) by which time the impact of the specified regime is expected to be adequately discernible in the household characteristics. At the same time, it will also be possible to crosscheck the simulation results with the 1981 Indian Census data (recall that our initial population is generated roughly approximating that of 1951 India).
It should be obvious by now that the model essentially works with three sets of specifications related to fertility, mortality and partitioning. Fertility has two components - nuptiality and marital fertility. To begin with, we consider two types of trend in both fertility and mortality - (i) static over the period of thirty years; and (ii) a decline over the period. This gives us four possible combinations of fertility and mortality trends and thereby four different demographic regimes. The parameter specifications are given in Table 4.
The regime corresponding to the decline in both fertility and mortality specified above should roughly approximate the demographic experience of rural India during the period 1951-81.

Each of the four demographic regimes was tried with two different specifications of partitioning probabilities. The corresponding mean critical household sizes were 10.0 and 13.5 respectively. Thus, we had eight sets of parametric specifications.

Two identical initial populations will not remain identical after a period of evolution under identical parametric regimes. A range of random variation occurs for a single fixed set of parameters. This point is particularly important when the size of the population is not large. In fact, the choice of parameters within a band of plausible values is not at all critical for this study but the variation that can occur for a given set of parameters needs to be taken care of.

Table 4: Specifications for Different Evolutions of the Initial Population

	Fertility				Mortality			
	Static		Decline		Static		Decline	
Period (Years)	Nuptiality	TMFR	Nuptiality	TMFR	e0(m)	e0(f)	e0(m)	e0(f)
1 - 5	India 1951-61	7.5	India 1951-61	7.5	36.0	35.0	36.0	35.0
6 - 10	India 1951-61	7.5	India 1951-61	7.0	36.0	35.0	40.0	39.0
11 - 15	India 1951-61	7.5	India 1961-71	7.0	36.0	35.0	43.0	42.[illegible]
16 - 20	India 1951-61	7.5	India 1961-71	6.5	36.0	35.0	46.0	45.5
21 - 25	India 1951-61	7.5	India(R)1971-81	6.0	36.0	35.0	50.0	45.5
26 - 30	India 1951-61	7.5	India(R)1971-81	5.5	36.0	35.0	53.0	53.0

For this purpose, evolution of the initial population over thirty years under each set of specifications was repeated four times to generate four independent samples. These four samples corresponding to each set of specifications will provide information about the nature of variations. We have taken averages of every relevant statistic (average household size etc.) across the these four samples while examining the results. The analysis of the results are presented and discussed in the next section.

4. Analysis of the Simulation Output

4.1 For the sake of brevity we shall refer to the two types of fertility and mortality specifications simply as 'static' and decline'. Similarly, we shall refer to the mean critical size specifications of 13.5 and 10.0 as 'low' and 'high' respectively (a low partitioning rate being associated with a longer period before splitting).

4.2 The average crude birth and death rates obtained in the simulation runs with high rate of partitioning are presented in Table 5 below. Columns (6) and (7) in the above table corresponds to the specification which is expected to simulate rural India for the period 1950-80. The CBR, CDR series resemble fairly closely that of the corresponding population estimates. The table of CBR, CDR corresponding to the high rate of partitioning look very similar, which is expected as the fertility and mortality parameters are the same in both the cases. It seems unnecessary to present the latter table here.

Table 5: Period-wise Average CBR and CDR Obtained in the Simulation Runs with High Rate of Partitioning

	Static Mortality				Declining Mortality			
	Declining Fertility		Static Fertility		Declining Fertility		Static Fertility	
Period	CBR	CDR	CBR	CDR	CBR	CDR	CBR	CDR
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ist 5-Year	42	27	42	27	42	27	42	27
2nd 5-Year	38	27	42	28	40	23	42	24
3rd 5-Year	40	27	43	28	40	21	42	22
4th 5-Year	39	27	42	27	39	19	42	19
5th 5-Year	38	26	43	27	35	16	42	17
6th 5-Year	35	25	43	27	33	13	41	14

4.3 The annual rates of growth (compound) of the number of households as well as that of the population are presented in Table 6 below.
A comparison of rows 1 and 2 with 3 and 4 (similarly, 5 & 6 with 7 & 8) seem to indicate that mortality differentials have a greater impact than that of fertility on the growth of the number of households. This observation should of course be qualified the range of parametric variation for the simulation runs.
As stated earlier, the third and the seventh rows corresponds to the simulation of rural India for the period 1950-80. The annual growth rates of rural population in India during corresponding decades (as given by the census

data) were - 1.88, 1.97, 1.75 respectively.

Table 6: Period-wise Average Compound Rate of Growth of Number of Households and Population Obtained in the Simulation Runs

Rate of Partitioning of Households	Mortality	Fertility	Ten-year Period Compound Rate of Growth (%)					
			Households			Population		
			First	Second	Third	First	Second	Third
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. High	Static	Decline	0.96	1.02	1.05	1.19	1.17	1.06
2. High	Static	Static	1.08	1.10	1.38	1.38	1.41	1.52
3. High	Decline	Decline	1.21	1.45	1.53	1.51	1.93	1.89
4. High	Decline	Static	1.41	1.71	2.01	1.56	2.18	2.55
5. Low	Static	Decline	0.19	0.71	0.80	1.28	1.16	1.05
6. Low	Static	Static	0.17	0.78	1.17	1.39	1.49	1.45
7. Low	Decline	Decline	0.26	1.09	1.64	1.42	1.95	1.93
8. Low	Decline	Static	0.37	1.19	1.71	1.67	2.06	2.50

4.4 Finally, let us examine the distribution of household size obtained under different parametric regimes specified. Table 7 presents the average household size and the tails of its distribution respectively.

Let us now consider columns (5) and (6) above which give the lower and upper tails of the distribution of household size. Column (5) gives the proportion of households of size less than or equal to 3. This proportion, for a given rate of partitioning, increases by about 2% if fertility declines without any decline in mortality, and falls by about 7% if mortality declines without any decline in fertility. Partitioning rate makes a difference within a range of 3 to 5% if fertility and mortality remain the same. However, partitioning rate makes a difference of 10% or more in the proportion of households of size 8 or above (column (6)). In summary, what seems to emerge is that differences in the rate of formation of new households create higher differentials in the average household size than differences in the fertility or mortality parameters. In other words, fertility and mortality differences make marginal differences, in a comparative sense, to the lower and upper tails of household size distribution and consequently to the average size whereas a lower rate of partitioning elongates the upper tail considerably leading to higher average size. As a result, differences in average household size are much narrower due to differences in demographic differentials than due to differentials in the rate of partitioning.

Table 7: Distribution of Household Si Obtained in the Simulation Runs

Rate of Partitioning of Households	Mortality	Fertility	Average Household Size	%Household of Size ≤ 3	%Household of Size ≥ 8
(1)	(2)	(3)	(4)	(5)	(6)
1. High	Static	Decline	5.02	27.0	14.0
2. High	Static	Static	5.21	25.2	16.9
3. High	Decline	Decline	5.37	21.5	17.3
4. High	Decline	Static	5.56	20.8	20.6
5. Low	Static	Decline	5.77	23.5	24.5
6. Low	Static	Static	6.01	21.6	28.5
7. Low	Decline	Decline	6.11	18.7	26.9
8. Low	Decline	Static	6.47	17.7	33.3

Note: All the figures above are computed averages across the four samples. Column (4) in Table 7 above seems to indicate that fertility and mortality differentials do not create as much differences in the average household size as the rate of partitioning does. Further, given a rate of partitioning, decline in mortality leads to larger differences in the average household size than decline in fertility.

Needless to say this observation is valid for rural India to the extent that the specifications for the simulation experiments cover the range actual variations in the concerned parameters.

4.5 One way of quantifying these relationships with the average household size discussed above is to obtain the relevant elasticities. With this purpose in mind we have attempted another set of simulation experiments. Instead of considering trends in the demographic rates we chose to study the dynamics of the initial population (over the 30 year period) with rates invariant over time. Rates were varied only from one simulation to another. Keeping the marriage probability fixed at India 1951-60 level, we considered three levels of TMFR - 7.5, 6.0 and 4.5. Similarly, we considered three levels of mortality as follows.

Specified Level	Life Expectancy at Birth $e_0(m)$	$e_0(m)$
1	53	53.0
2	46	45.5
3	36	35.0

These give us nine combinations fertility and mortality levels. Simulated populations were evolved separately with each of these nine sets of parameters at two levels of mean critical household size (as before) - 10.0 and 13.5. The number of replications for each set of specifications was limited to two this time. Thus, we designed the experiment to produce 36 (9 x 2 x 2) observations for estimating the concerned relationships in terms of relevant elasticities.

4.6 It should be easy to see that as a result of the above simulation experiments we obtained 36 independent sample observations - 2 corresponding to each set of given parameters $e_0(m)$, $e_0(f)$, TMFR and mean critical household size. There were 18 such sets of parameters. Now, the relevant elasticities could be estimated by regressing average household size obtained on the corresponding pre-specified set of parameter values. The regression results are presented in Table 8 below.

Table 8: Regression of Average Household Size on Mean Critical Household Size, Female Life Expectancy at Birth and Total Marital Fertility Rate

Note: All <u>variables</u> considered under natural <u>logarithmic</u> transformation

<u>Dependent Variable</u>: Average Household Size

Independent Variables	Regression Co-efficient	Standardised Co-efficient	T-Statistics
1. Mean Critical Household Size	0.44418	0.66209	9.547
2. Female Life Expectancy at Birth	0.25692	0.43736	6.306
3. Total Marital Fertility	0.22398	0.46524	6.708

R-Square 0.85 Standard Error 0.04189 F-Statistics 58.64

Note: All test statistics were significant at 1% level

The SPSS/PC programme was used for the purpose of estimating the regression with Method=Enter. The programme excluded the variable Male Life Expectancy for which the relevant F-value was found to be insignificant. This is understandable since by our specification the two life expectancy variables were nearly collinear and so one of them would have to be excluded naturally. Since the variables in the regression were used after natural logarithmic transformation the estimated regression co-efficients are the corresponding estimated elasticities. It seems quite obvious (see Table 8) that the impact of partitioning rate (measured by the mean critical household size) on average household size is considerably higher than that of fertility or mortality. Further, it is interesting to note that the sensitivity of average household size to mortality is a little higher than that to fertility - the elasticities are 0.26 and 0.22 respectively (which could of course be in part due to sampling variation).

5. Relevance to Landholdings Data

5.1 We have referred in the introduction to data on household size variations in rural India as the point of departure for these simulations. These data exhibit a considerable variation in the mean size corresponding to the different landholding classes - ranging from 4 or below in the lowest class to 8 or above in the largest class.

Now, for any given specification of demographic rates, a distribution of households by size emerges out of random variation with small and big families occurring with certain frequencies depending on the parameters. However, the occurrence of large differences in the mean size in fairly large (and possibly, relatively homogenous) population groups such as those of large landowners and agricultural labourers, would still require an analysis and explanation, for, variations in the mean size have to be considerably narrower than in distributions of individual families.

5.2 It is in this context that we have set up the simulation model discussed in this paper. What the results show convincingly is that given demographic variation in the range actually observed in rural India over the past 4 or 5 decades, the mean household size of any large subgroup of the population is likely to be in range of 4 to 6 unless the partitioning rate for the sub group is either too low (leading to a high mean size such as 8) or too high (leading to a low mean size below 4). In particular the simulations suggest that the large mean size associated with the class of big landowners - a consistent feature over the entire period 19[illegible]-81, as revealed by the different landholding surveys covering this period - must have clearly arisen as a consequence of low splitting rate, i.e., a greater propensity for large landowning households to remain 'joint' in some form or the other, in comparison to small landowners and agricultural labourers among whom the proportion of 'nuclear' households tends to be very high.

5.3 Of course, the differences in demographic parameters in between the different landowning classes may also contribute in some measure to the observed variation in mean household size: thus, for example, among the poorer classes mortality rates could be higher and fertility rates lower. In the absence of the relevant data the elasticities we have derived here through simulation would be of some help in assessing the relative contributions of demographic parameters on the one hand and the patterns of household formation on the other.

APPENDIX

Design of the Modules

Descriptions of all the modules in the computer simulation model are presented here along with the nature of inputs and some of the outputs for validation.

A Marriage Module

A.1 First marriage probabilities are applied to males of age between 10 and 54 and to females of age between 5 to 49 (inclusive of the limits). We have used the nuptiality rates (proportion single) as estimated from decade synthetic cohorts based on the Indian Census data for specifying the marriage probabilities of never-married persons. The annual probabilities of marriage are computed from the proportion of never married in each five-year age group under the assumption of a geometric distribution. It should be noted here that the marriage probabilities are being specified on the basis of macro level data on nuptiality rates for microsimulation. Since our primary concern is to generate an appropriate distribution of age at marriage this procedure (for specification) should be acceptable. The difference between the age of the husband and that of the wife has been kept constant at 5 years. Only the male marriage probability has been used for the occurrence of marriage. In other words, the annual probability of a female getting married has been taken to be the same as that of a male 5 years older than her. The simulation model is an open one so far as marriage is concerned. Females marry someone outside the population and thus drop out of the population list, and males marry from outside the population and the wife gets included into the list (which is realistic at the household level). The marriage probability, the way it is applied here, thus tends to keep the balance between females dropping out and coming in through marriage.

A.2 No divorce or separation is allowed. It was found rather difficult to empirically specify the remarriage probabilities. The necessary data (Indian) for this purpose are not readily available. We have introduced certain probabilities of remarriages based on the estimates provided by Bhat (1984)[6]. According to Bhat's estimate 62.44% of the widowers are found to have remarried in the 1981 census, and the figure is 33.68 % for females. What has been specified in the model is as follows -

A widower remarries with probability 0.6244xp(a) here p(a) is the probability of marriage of a never-married male at age 'a'. A widow with children does not remarry. A widow without children remarries with 50% of the probability of the corresponding never-married female.

A.3 A few of the average ages at marriage obtained in the course of various simulation runs are presented Table 9.

Table 9: Average Age at Marriage obtained in some Simulation Runs

Specification	Average Age at Marriage			
	Estimated Age at Marriage		Obtained in Simulation	
	Male	Female	Male	Female
India 1941-51	19.93	15.43	21.06	16.37
India 1951-61	22.32	16.10	21.99	16.93
India 1961-71	22.72	17.12	23.11	18.35
India (Rural) 1971-81	23.66	18.50	24.36	19.14

Note: The estimated average ages at marriage were based on decade synthetic cohort and reported in the ESCAP Monograph (1982)[7] for three decades beginning 1941. The figures for 1971-81 were estimated by the authors following the same procedure.

The average age obtained in the simulation runs are higher than those estimated for the corresponding decades by about one year in most cases. This is due to the fact that census estimates are based on nuptiality rates for the population in the age group 0 to 50 years whereas in the simulation model males marry between age 10 to 54 and the brides age is set to be 5 years below the groom's age.

B Birth Module

B.1 Child birth probability depends on the woman's age, marital status and parity i.e. total number of children she has already borne. The marital fertility curve is specified by a function b(x,y,z) as given below. Notice that there is no explicit specification for birth spacing except a very simple assumption that a married woman can not bear more than one child during a period two of successive years.

Marital Fertility Function

$$b(age, lage, prt) = 0 \qquad \text{if } age=lage+1$$

$$= \exp(aa + bb.age + cc.age^2 + (1-prt).pf) \qquad \text{otherwise}$$

where age= current age of the female
lage= age at which the last child was born
prt= total no. of children born to the female so far (current parity)
aa, bb, cc, pf are given constants.

B.2 The parameters aa, bb, cc and pf have been empirically specified with some trial runs to reproduce the marital fertility profile more or less similar to what is observed in the Indian case. Initially, aa through cc were estimated by regressing proportion of married females giving birth to the first child in different age intervals as recorded by Census of India (1981). A semi-log quadratic function was used for the purpose. In the next step, the constant aa was adjusted through a few trial simulation runs to obtain Total Marital Fertility Rate (TMFR) about 8.5, keeping pf set to zero. The TMFR specification can now be adjusted by choosing suitable values for pf. These values were determined by simulation runs again.

Figure 1

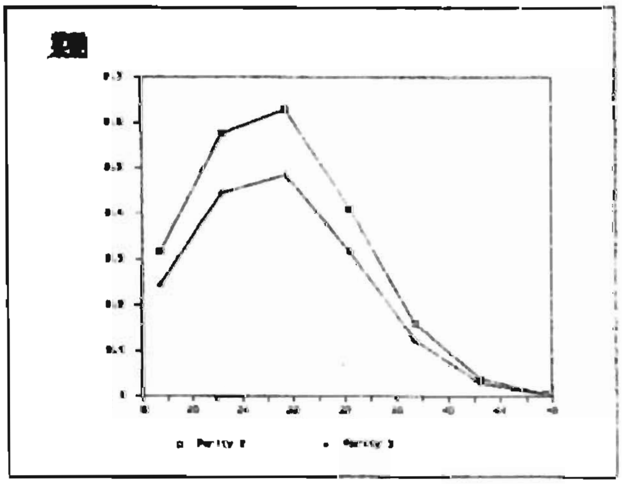


Figure 1 shows the marital fertility curves for the second and third child birth described by the chosen function when TMFR is set to 5.0.

The gender of the child born was determined by a binomial event with probability for the child to be male set to produce the sex ratio (male/femalex100) at birth to be 106 (roughly corresponding to the 1971 all-India figure).

Age-specific average parities of currently married women who married between age 15-19 years as obtained in the simulation year 161 while generating the initial population with TMFR specified at 7.5 are presented below. Note that simulation year 161 corresponds year 1951 in our design. For comparison we have taken estimates of average parities of currently married women whose age at first child birth was between age 15-19 years, see Table 10. These estimates are based on a 10% sampl in the Travanco -Cochin region (at present part of the state of Kerala, India) during 1951 Census.[9] For comparison we have taken estimates of average parities of currently married women whose age at first child birth was between age 15-19 years, see Table 10. These estimates are based on a 10% sample in the Travancore-Cochin region (at present part of the state of Kerala, India) during 1951 Census.[9]

2.24		
25 - 29	3.3	4.25
30 - 34	4.8	5.91
35 - 39	6.0	6.60
40 - 44	6.8	6.94
45 +	7.6	7.34

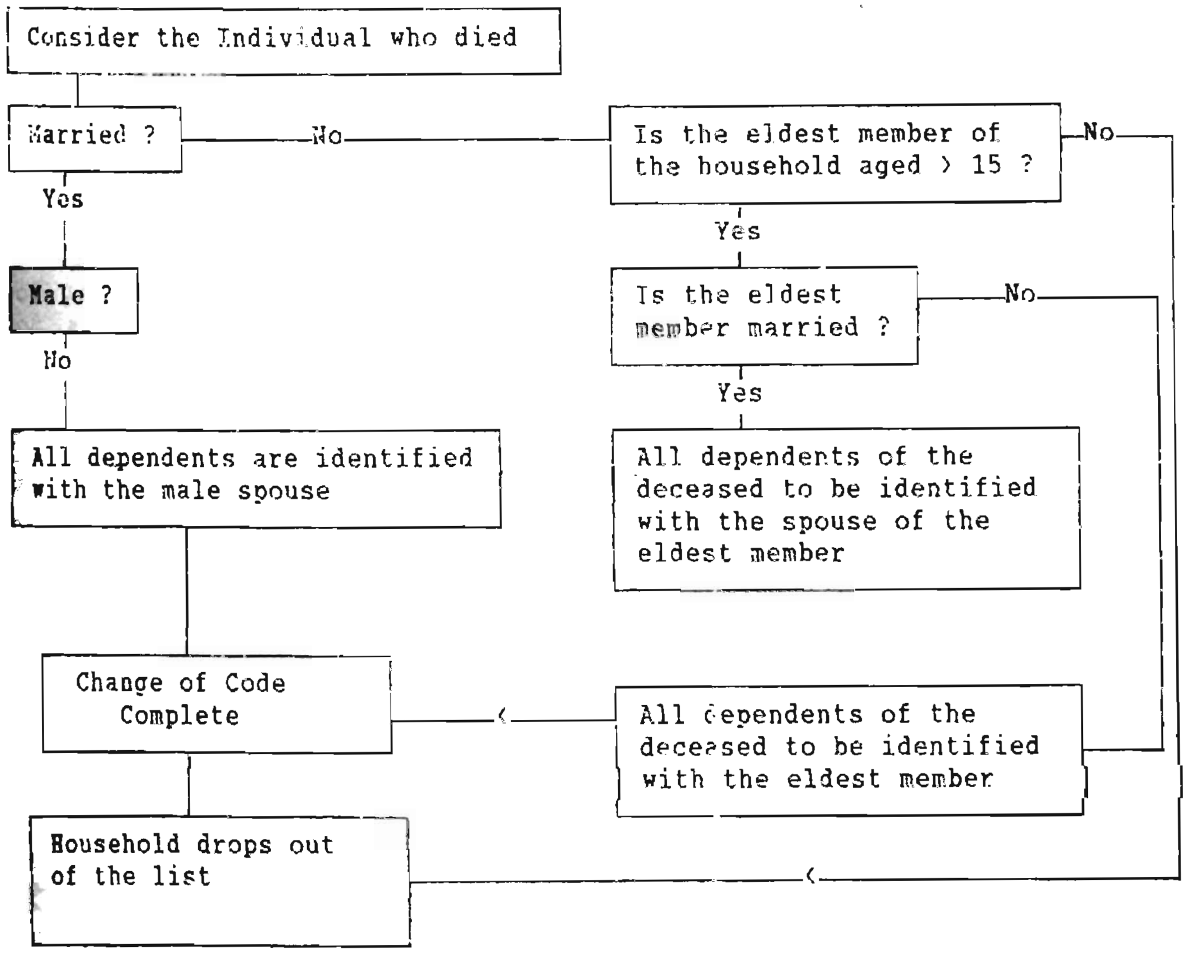
* Graduated Estimate based on 10% sample, currently married women with age at first childbirth between age 15-19 years

C Death Module

C.1 Coale-Demeny Regional Life Tables were used to specify the death probabilities after deciding upon the life expectancy at birth for a given year. Single age-specific annual death probabilities were computed on the basis of an assumption of geometric distribution of the event over the 5 year age interval.

C.2 When a death occurs there is a follow up action which relates to changes in the 'dependency code'. As stated in section 2 of the text the dependency code essentially keeps track of relationships among the individuals within a household in a simplified form.

This is used for the purpose of partitioning the household as and when a household breaks up. The chart above explains the rule for change of the code in the event of a death. Notice that if there be no one aged above 15 years left in the household after the death of an individual the household is dropped out of the list i.e. we assume that the children will join some other household.



C.3 We have not computed any statistics on life expectancy from the simulation output. However, the crude death and birth rates obtained in the simulation runs have been presented later in this Appendix.

D <u>Formation of New Households</u>

D.1 This section of the model is based on a specification of the probability of splitting, which is a function of the size of the parent household and a convention for partitioning the household to form new households. The probability is set to zero for households with less than two living couples. The probability that a given household (with at least two living couples) will break up is given by a function p(x) described below.

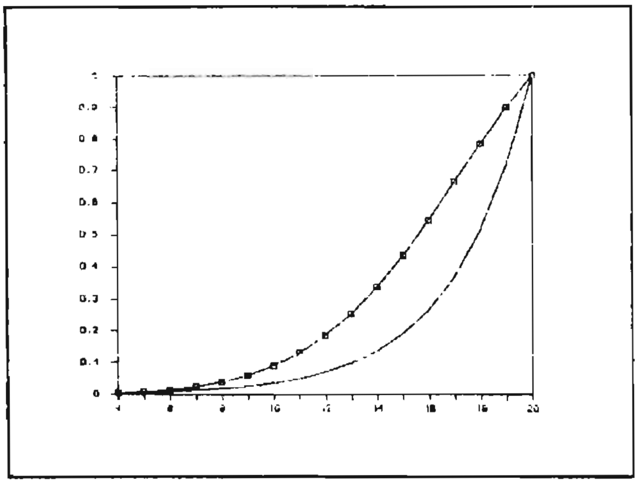
$$p(s) = 0 \quad \text{if number of couples in the household} < 2$$
$$= \min\,[apt.\exp(bpt.ss+cpt.ss^2)\ ,\ 1\ ]$$

where

s = household size
$ss = s-4$
$apt = 0.01$
$bpt = -\ (\ln(apt)+256.cpt)/16$

cpt is a given constant the value of which is adjusted to obtain slow or fast rate of formation of new households.

Figure 2



The function has been arbitrarily decided upon after some trial and error. The specification is such that the probability is unity for s=20 i.e. household size is not allowed to exceed 20. This function implies that the chance of a household breaking up increases with it's size and it reaches unity at or before the size is 20. Figure 2 provides an example of probability specification using this function. The marked curve corresponds to cpt=-0.015 and the other one for cpt=0.00.

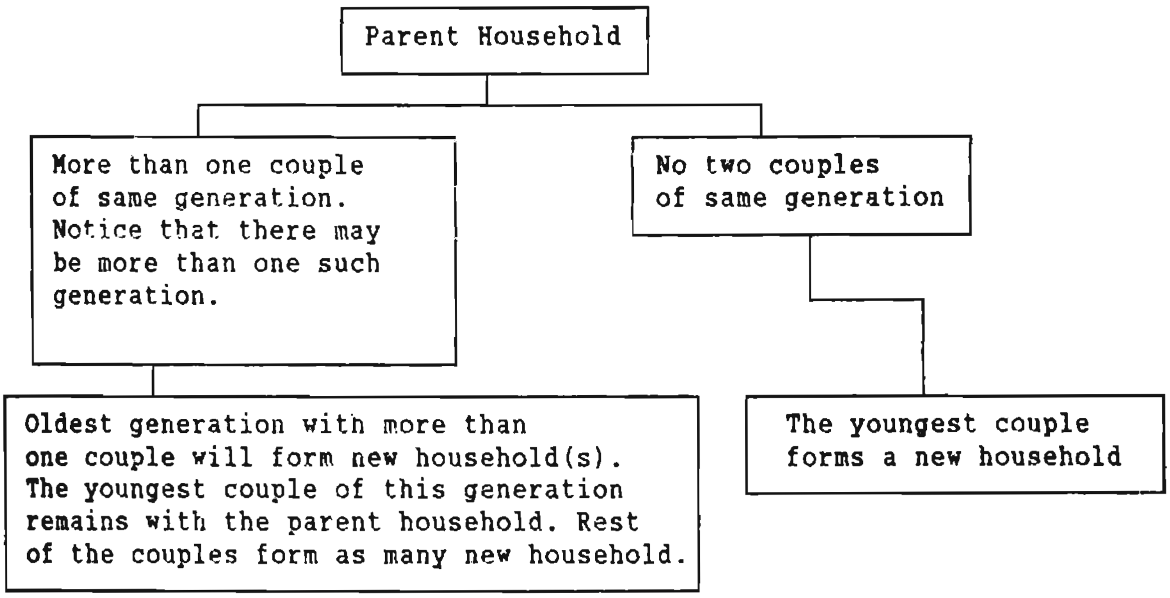
D.2 Let us define 'Critical Size' as the size of a household when it breaks up or partitions. Given the above probability specification one can compute the Mean Critical Size for a given value of the parameter cpt. It is the average size of the households at the time of their partitioning. A specified value of cpt does not give one any impression of the implied rate of partitioning. We shall, in our discussion, use the mean critical size corresponding to a given cpt as a measure of specified nature of partitioning.

D.3 When a household breaks up according to the above probability specification, it is partitioned (i.e. forming new households) according to a rule as described below.

The rule for splitting the household is based on the village study reported by Shah(1974)[10]. According to this report - most often households partition only after sons get married and have children. It was also found that it is the youngest married son who stays back with the parent household in almost every case of partitioning. Although these observations are based on only a single village study in Gujarat and the partitioning conventions are likely to vary from region to region, we have based our convention of partitioning a household in the model on these broad findings.

D.4 We refer to existence of two or more couples of the same generation in a household as 'jointness', and two or more couples of different generations as 'extension'. Notice that it is possible, at least theoretically, to have jointness in more than one generation within a household. For example, married brothers and their married children may live in the same household. This household, in our terms, is both joint and extended.

D.5 The principle that we follow is to partition a household to break the jointness of the eldest generation, if it exists. If there be no jointness in any of the generations within the household then the youngest couple forms a new household. In the case of partitioning the joint household, the youngest couple of the generation remains in the parent household, other couples of the generation forms as many new households, each along with their dependents (according to the dependency code). The following chart illustrates the principle.



E Overall Performance of the Model

E.1 We have shown above that the marriage and the birth module perform reasonably well to produce acceptable average age at first marriage and marital fertility profile. Overall performance of the model in terms of mortality, fertility and nuptiality, however, can be assessed only by examining the resultant age and marital status distributions. For this purpose, we consider the initial population generated. Recall that this population is generated to approximately resemble that of Rural India 1951.

E.2 A comparison of the age-sex distribution of the initial population generated by the model and what was obtained in the 1951 Census of India (Rural) is presented Table 11. The Census distribution is estimated from a 10% sample of the population. One should also note that the unit of time in the simulation is one year i.e. age is recorded in integral years. Thus, the two age distributions are not strictly comparable. However, the closeness between the two distributions is quite satisfactory in terms of the fertility-mortality performance of the model.

Table 11 : Comparison of Sex-wise Age Distribution of the Simulated Population and Census of India (Rural) 1951

Age Group (in years)	Male		Female	
	Simulation	Census 1951	Simulation	Census 1951
0 - 4	14.7	13.5	13.8	13.9
5 - 14	24.6	25.4	23.4	24.7
15 - 24	17.7	16.5	17.9	17.3
25 - 34	14.6	15.2	14.3	15.7
35 - 44	12.9	12.2	11.7	11.6
45 - 54	7.6	[illegible].8	8.8	8.2
55 - 64	4.7	5.2	6.5	5.1
65 & above	3.2	3.2	3.6	3.5
All	100.0	100.0	100.0	100.0

E.3 Table 12 compares the distribution of marital status of the simulated initial population with the Indian census figures of 1951. Proportions of unmarried and widowed population obtained in the simulation are higher than the corresponding census figures.

Table 12: Comparison of Distributions of Marital Status of the Simulated Population and Census of India 1951

Marital Status	Male		Female	
	Simulation	Census 1951	Simulation	Census 1951
Unmarried	52.8	48.8	42.0	38.2
Married	38.7	46.0	41.4	49.0
Widowed	8.5	5.2	16.6	12.8
All	100.0	100.0	100.0	100.0

Higher proportion (about 4% in excess) of unmarried persons is due to the same reason which causes higher average age at first marriage noted in section A.3 above. Higher proportion of widowed population is most probably due to the specification of remarriage probabilities. Specification of high death probabilities could have led to the higher incidence of widow/widowerhood but the crude death rate obtained in the simulation (presented in section 4.2, Tablethe 5) seems to be quite satisfactory. Hence we rule out this possibility.

F FAMSIM as Compared to SOCSIM

F.1 The major simplification in FAMSIM begins with the unit of time being an year instead of a month, as it is in the case of SOCSIM. Apart from reducing the model run time on the computer this implies simplification of several features of the model. For example, in SOCSIM, infant mortality is specified in two phases - the first month of life and the next 11 months, whereas in FAMSIM it is only an annual event.

F.2 The Birth module of our model is probably the most simplified one compared to SOCSIM. A random birth spacing routine in SOCSIM adjusts a woman's probability of giving birth soon after bearing a child depending on whether the child survives or dies. Childbirth probabilities are modified by a fertility multiplier assigned at random to the woman at birth and carried with her through life, making some women consistently subfertile and some superfertile. The latter provision is important as this heterogeneity in fertility ensures the variation in completed family size observed in reality. However, some of these features can be easily incorporated into FAMSIM at a later date.

REFERENCES

1. Krishnaji, N (1980): 'Agrarian Structure and Family Formation - A Tentative Hypothesis', Economic and Political Weekly, Review of Agriculture, Vol 15, No 13, pp. A38-A43.

2. Laslett, Peter (1972): Household and Family in Past Time, Cambridge University Press.

3. Burch, Thomas K. (1972): 'Some Determinants of Average Household Size: An Analytical Approach', in Laslett(ed) (1972).

4. Watchter, K.W. et.al. (1978): Statistical Studies of Historical Social Structure, Academic Press, New York.

5. Hammel, E.A. et al. (1976): The SOCSIM Demographic-Sociological Microsimulation Program: Operating Manual, Research Series no. 27, Institute of International Studies, University of California, Berkeley.

6. Bhat, M. and Kanbargi, R. (1984): 'Estimating Incidence of Widow and Widower Re-marriages in India from Census Data', Population Studies, March 1984.

7. Population of India, Country Monograph Series No. 10, Economic and Social Commission for Asia and The Pacific, Bangkok, Thailand, 1982.

8. 'Estimation of Birth and Death Rates in India during 1941-50', Census of India 1951, Paper No. 6, 1954, Registrards year 1951 in our desig

9. 'Estimation of Birth and Death Rates in India during 1941-50', Census of India 1951, Paper No. 6, 1954, Registrar
20 - 24

10. Shah, A.M. (1974), The Household Dimension Of the Family in India, University of California Press, Berkeley, 1974.