Fertility Transition and Threshold Estimation: A District - Level Analysis in India

Abhiman Das*

Abstract

This article focusses on issues relating to fertility transition and related socio-economic variables. The observed differential in fertility between different states, as determined by the cluster-cum-discriminant analysis, by using district-level data, clearly establishes the link between fertility change and social backwardness of women, especially in respect of female education and age at marriage. The economic variables, on the other hand, are found to be less important for the existing fertility differential between states. The findings suggest that the threshold of female literacy for a faster fertility decline in India is about 43 per cent; once that level is achieved, fertility rate will decline faster towards the stability of the population.

Introduction

The demographic diversity of India poses a challenge to planners and policy makers. Physical accessibility to scarce resources depends on several factors, including population distribution. Geographers always classify regions and sub-regions according to factors like soil, rainfall, mountains, river basins. Economists, on the other hand, have attempted to classify regions on the basis of economic criteria, but because of the data constraint, their work has not been as detailed as that of geographers (Bose 1994). A rigorous and pioneering contribution in this direction in India can be said to have originated with the work of Mitra (1965), who attempted to classify all the districts in India on the basis of a large number of variables for which census data were available. Of late, there have been several studies on economic regionalisation. A diagnostic regional analysis of the shortfalls in development and utilisation of human resources by Pathak (1991) has made a significant contribution towards the spatial variations of districts in India using

* Assistant Adviser, Reserve Bank of India, Department of Statistical Analysis & Computer Services, DBS, C-9, 7th Floor, B-K Complex, Bandra (E), Mumbai – 400 051.
E-mail: adas@rbi.org.in; abhiman_das@yahoo.com.

I sincerely acknowledge the academic help received from my Ph.D. Supervisor Professor K. B. Pathak, Director (Retd.), International Institute of Population Sciences, Mumbai, in finalising this paper. I am thankful to Dr. R. B. Barman, Officer-in-Charge, Department of Statistical Analysis and Computer Services, Reserve Bank of India, Mumbai, for his insightful thoughts on some of the analytical issues relating to social science research.

census data. It may be mentioned that broad state-level comparisons and classifications may not be able to capture fully the extent of diversities among various indicators characterising several facets of development. Nevertheless, state-level indicators are of prime importance as the state is a crucial political unit. A wide range of relevant fields, including health and education, are constitutionally defined as ‘state subjects’, to be handled at the level of the states rather than of the central government. There are also ‘concurrent subjects’, involving both state and central governments (Sen and Dreze 1998).

Considerable regional diversity in terms of social, economic and demographic characteristics prevails in India. This is true not only among the states but also among the districts of the same state. In general, these striking variations among the states in the livelihood of the common people stem from various factors such as the level of literacy, female education, nutritional standards, infant mortality, morbidity, employment, income distribution, public distribution system, political commitments, etc., and their corresponding interactions. Thus, any country-level study is likely to hide variations at the micro level. ‘The Indian subcontinent, with its large size, wide structure and eco-social disparities, is better understood and better interpreted when studied at the regional level. Analysis of data in disaggregated form narrows down the variability and enables better identification of special characteristics’ (Datta Roy Choudhury 1995). The question of classification of the states into homogeneous groups acquires special status to initiate action programmes to bridge the gap among states (Guru 1992). Different states are at different stages of demographic transition. States differ greatly in respect of mortality decline too (Navaneetham 1993). The factors responsible for demographic changes have been different for different classes of states. Since demographic parameters reflect the progress and development of society, they can be safely used to classify the states or districts in different classes associated with different stages of transition. The set of variables characterising the class constitutes an important link in formulating action programmes. The relative importance of the variables and the threshold at which they become significant in affecting development varies over time. Again, the time for demographic change, varying from class to class, is determined by the interaction of socio-economic and psychological factors.

This paper proposes to explore the main reasons for the existing imbalances using Bayesian discriminant analysis, which can accommodate the prior information of the variables for classifying a state, using the district as a unit of observation. The specific objectives are: 1) to classify major Indian states into some homogeneous groups based on the level of demographic transition and to derive a set of linear discriminant functions to indicate the group to which a state belongs; 2) to determine a lower bound and a threshold value of a particular variable of a state required for a shift to a specific demographic transition phase; 3) to estimate the time required for a state to shift from a lower transition phase class to an upper transition class and 4) to study the peculiarities of the districts that are subsequently misclassified and show some potential shift of these districts.

This study would be useful for predicting the levels of different inputs necessary for setting up plans or programmes and to prepare a sketch of the time frame of transition from the policy point of view. With the approach adopted here, it will be easy to locate misclassified districts and prepare their individual case studies. Also, the use of the Bayesian approach instead of the traditional classical approach in discriminant analysis helps one to unravel the change occurring in the class over time.

Methodology

The first sub-section, 'Classification', outlines the methodology for classifying major states into homogeneous groups, while the second sub-section, 'Discrimination', is devoted to the methodology required for Bayesian discrimination of the districts.

Classification

Fertility behaviour is a complex phenomenon that results from the interplay of various social, economic, psychological and cultural patterns related to marriage, childbirth, child rearing and kinship affiliation. It is not feasible to explain all the factors of fertility transition in India together. Therefore, we need to classify states into groups in such a way that they are highly heterogeneous between groups and homogeneous within groups to better understand diversity in their fertility transition.

To avoid subjectivity and classification bias, cluster analysis (CA) has been used to classify sixteen major states. The general aim of CA is to 'allocate a set of individuals to a set of mutually exclusive, exhaustive groups such that individuals within a group are similar to one another while individuals in different groups are dissimilar' (Chatfield and Collins 1980). However, there is no completely satisfactory way of defining a cluster. A number of clustering methods are available, each of which will often produce structures that are substantially different. This is because the choice of a clustering method implicitly imposes a structure on the population and is often tantamount to defining a cluster. If a classification does exist, a further problem is that the data may admit more than one different but meaningful classification depending on the purpose of the investigation. Another point to bear in mind is the possible existence of proper information on the likely number of groups. This will be helpful in finding a partition. However, Anderberg (1973) points out that this prior information could be misleading if the data are sampled from larger population and one (or more) of the groups happen to have been excluded from the sample. If the analyst still tries to find the full number of groups, 'silly' clusters may be created. Bearing in mind the wide variety of practical situations, it is rather hard to make general recommendations. Therefore, instead of imposition of a number of clusters from *a priori* knowledge, the graphical approaches of clustering have become more popular. The more sophisticated graphical approach, known as 'dendrogram' analysis, has been used considerably. But there are also a number of methods available for sketching a dendrogram. In this study, the 'average linkage' method has been used to classify the states. The distance between two clusters is the average distance between pairs of observations, one in each cluster. Average linkage tends to join clusters with small variance and is slightly biased toward producing clusters with the same variance.

Discrimination

The basic idea of discriminant analysis is to allocate an observation to one of the classes specified earlier on the basis of a discriminant function derived from a whole set of observations from each class.

Suppose we have 'g' groups denoted by π_i, i =1, 2, ………..g.

Let $f_i(x)$ be the density associated with population π_i, i =1, 2, ………..g. Let p_i be the prior probability of population π_i, i =1, 2, ………..g. C(k|i) = the cost of allocating an item to π_k, when in fact it belongs to π_i, for i ($\neq$k) = 1, 2, ………..g.

For k=i, C(i|i) = 0. Let R_k be the set of X's classified as π_k

and prob.(k|i) = prob. (classify item $\pi_k|\pi_i$) $= \int_{R_k} f_i(\mathbf{x})\, d\mathbf{x}$ for k, i = 1, 2, 3, ……………, g

with $p(i|i) = 1 - \sum_{k\neq i,\, k=1} \text{prob.}(k|i)$

Thus, expected cost of misclassification

$$(ECM) = \sum_{i=1}^{g} p_i \{\sum_{k\neq i\, k=1}^{g} p(k|i)\, C(k|i), \ldots\ldots(1)$$

The classification regions that minimize the ECM are defined by allocating $\mathbf{x}$ to that population π_k, k=1, 2,….., g

for which $\sum_{i=1\ \ k\neq i}^{g} p_i f_i(\mathbf{x})\, C(k|i)$ is smallest.

Assuming (0-1) loss function, it can be proved that [Anderson (1984), Johnson and Wichern (1996)] the minimum ECM classification rule is: Allocate $\mathbf{x}$ to π_k if $\ln p_k f_k(\mathbf{x}) > \ln p_i f_i(\mathbf{x})$ for all i$\neq$k.

We assume that the distributions of variables in all the groups follow multivariate normal distribution with mean μ_i and variance-covariance matrix Σ_i, i= 1, 2, …………,g.

Therefore, allocate $\mathbf{x}$ to π_k if

$\ln p_k f_k(\mathbf{x}) = \ln p_k - (p/2)\ln(2\pi) - (½) \ln |\Sigma_k| - (½) (\mathbf{x} - \mu)' \Sigma_k^{-1} (\mathbf{x} - \mu) = \max. \ln p_i f_i(\mathbf{x})$, for all i.

The constant (p/2)ln(2π) is ignored as it is the same for all the populations. We define the discriminant score for the ith population to be $d_i(x) = \ln p_i - (½) \ln |\Sigma_i| - (½) (\mathbf{x} - \mu_i)^{/} \Sigma_i^{-1} (\mathbf{x} - \mu_i)$, i = 1, 2, …, g and the allocation rule is: Allocate **x** to π_k if $d_k(\mathbf{x})$ = largest of ($d_1(\mathbf{x}), d_2(\mathbf{x}), d_3(\mathbf{x}), \ldots\ldots, d_g(\mathbf{x})$). We further assume that the population covariance matrices, Σ_i for all i, are equal. When $\Sigma_i = \Sigma$ for all i, then $d_i(x) = -(½) \ln |\Sigma_i| - (½)\, \mathbf{x}^{/} \Sigma^{-1} \mathbf{x} + \mu_i^{/} \Sigma^{-1} \mathbf{x} - (½)\mu_i^{/} \Sigma^{-1}\mu_i + \ln p_i$. As the first two terms are the same for all groups, we can ignore them for allocatory purpose. Thus, by defining the linear discriminant function as $d_i(x) = \mu_i^{/} \Sigma^{-1} \mathbf{x} - (½)\mu_i^{/} \Sigma^{-1}\mu_i + \ln p_i$, the allocation rule is as follows: Allocate **x** to π_k if the linear discriminant score $d_k(\mathbf{x})$ = largest of ($d_1(\mathbf{x})$, $d_2(\mathbf{x}), d_3(\mathbf{x}), \ldots\ldots, d_g(\mathbf{x})$).

The steps are as follows: a) Sixteen major states are grouped into some classes from the observed dendrogram of usual cluster analysis based on some variables indicating different phases of fertility transition, b) A set of variables has been selected which can be considered as direct/indirect determinant of fertility behaviour by taking the district as a unit of observation, *i.e*., characteristics of districts are as if samples of the respective states. Then a set of discriminant functions has been derived on the basis of selected variables; prior probabilities have been taken as the corresponding proportion of district in each group to the total. These priors indirectly put a weightage pattern for the discrimination. c) An upper bound of female literacy required for fertility transition for some selected states has been calculated by using the Bayesian allocation rule of discriminant procedure. d) a time frame required for fertility transition is recommended for some states with the assumption that female literacy in India follows a logistic law over time.

Selection of Variables

The districts in all states have undergone several important changes since independence. Some districts have witnessed rapid transformation, while the changes in others are quite slow. Virtually all socio-economic indicators are moving in the same direction of overall development; the expansion of educational opportunities, and the increase in the levels of literacy, enrolment and educational attainment; the improvement in health services and public sanitation, and the resulting decline in mortality; the transformation of the economy from an almost exclusively agricultural base to a mix of industry and agriculture; the spread of urbanisation; the electrification of many rural areas; the expansion of transport networks; and the rapid increase in the proportion of population reached by the mass media. But such social and economic changes have spread unevenly among regions and different segments of the population. It is often argued that economic liberalisation is not enough for development unless accompanied by social development. The high concentration of power and privileges deriving from the combined effects of inequalities based on class, caste and gender have created an environment extremely hostile to social change (Sen 1997). These have been reflected in the diversity of the reproductive patterns in the states/districts of India. Six variables are selected on the basis of two criteria: a) proximity of the variables to reflect the fertility behaviour of women and b) availability of district-level data. Six major broad components of variables have been considered here: (1) district income, (2) female education, (3) urbanisation, (4) employment in services sector, (5) per capita income and (6) age at marriage. All these components are measured at the district level.

District Income

Economic inequality between districts depends upon several factors such as agricultural output, industrial activities, employment, educational opportunities, level of urbanisation, geographical location, and transport facilities. But at the district level it is extremely difficult to measure these factors exactly, and all the factors are not equally important. Given the lack of availability of data, the only way out is to develop a composite index of the relative development of the district as a proxy for income indicators. The weighting pattern to develop this index should reflect, by and large, the importance of the different sectors of the economy. The weighting pattern, as devised by the CMIE (Centre for Monitoring Indian Economy), is as follows. The agriculture sector has a total weight of 35 per cent in the index. It includes the per capita value of output of crop (25 per cent) and per capita bank credit to agriculture (10 per cent). The mining and manufacturing sector is assigned a total weight of 25 per cent. Mining, manufacturing non-household and household workers per lakh population is given a weightage of 15 per cent. The per capita bank credit to industry is given 10 per cent weightage. The service sector carries 40 per cent weightage; per capita bank deposit (15 per cent), per capita bank credit to services (15 per cent), literacy (4 per cent) and urbanisation (6 per cent). The weighting pattern of CMIE is rather subjective. Therefore, the relative development index (RDEV) of a district thus developed can be used only as an approximate indicator for understanding the regional imbalances with respect to overall development between districts. The data for district-level RDEV is culled out of the CMIE publication on districts' profiles of India.

Education

Studies of fertility conditions and change have consistently pointed to education as an important factor in accounting for fertility differences within population. The specific connections that have been theorised to exist between education and fertility can be classified in various ways. To adopt the economist's terminology, individual-level effects of education can be divided into those that act on the demand for children, those that affect the supply of children and those that influence the costs, broadly defined, of fertility regulation (Easterlin 1978). Education reduces the expected long-run desire of wealth flow from children to parents (Caldwell 1983).

The connection between education and desired family size is well-established (United Nations 1987). Education may directly change attitudes, values and beliefs towards a small family norm and towards a style of child-rearing that is relatively costly to the parents in time and money. The potential for education to diffuse non-traditional values does not end in the classroom, since the educated are likely to continue to be exposed to modern ideas. Education also influences economic factors in ways that are thought to discourage high fertility; it reduces the economic utility of children; it creates aspirations for upward mobility and accumulation of wealth; it increases the opportunity cost of women's time and enhances the likelihood of their employment outside the home.

There is a strong positive relationship between education and contraceptive use too. Some of the avenues through which education may affect fertility control are: (a) by facilitating the acquisition of information about family planning; (b) by increasing husband-wife communication; (c) by imparting a sense of control over one's destiny, which may encourage attempts to control childbearing as well; (d) the higher income group of educated couples makes a wide range of contraceptive affordable. Education also affects the supply of living children through paths other than its influence on deliberate fertility control. The two most important of these influences are: (a) education delays entry into marital unions; (b) education is associated with reduced child and adult mortality. The variable chosen from this component is female literacy rate (FLIT) and the corresponding data are taken from the 1991 Census.

Urbanisation

Urbanisation has been used as a proxy for modernisation in studying fertility behaviour as it helps in modifying the natural fertility of married women through increased use of contraception. Also, it affects age at marriage. It helps to enhance the status of women in society by improving literacy and educational levels, increases their involvement in productive employment outside the household sector and enhances the longevity and health conditions because of better facilities and awareness of these, and hence leads to better utilisation of maternal and child health care services. In societies undergoing social and economic transition due to development, urban living and lifestyles promote the desire for smaller families and hence reduction in fertility. In short, urbanisation affects fertility through characteristics such as availability of educational opportunities, health facilities, job opportunities in the modern sector, communication facilities and contraceptive information and supplies; and the cost of fertility regulation and of bearing and rearing children. The data on level of urbanisation (URBAN) are taken from the 1991 Census.

Employment

Increased labour force participation, especially of women, has been proposed repeatedly in both the demographic literature and population policy statements as a means of promoting development and reducing fertility in developing countries. A great deal of empirical work has been carried out to examine the connection between employment and fertility of women in developing countries (WFS, United Nations 1987). The hypothesis that women's employment is negatively related to fertility receives support from most empirical studies. Women's employment has seven specific indirect effects in social life — maternal, conjugal, domestic, occupational, kin, community and individual-cum-psychological, each having an impact on the opportunity costs of children and hence on fertility (Oppong 1983). Additional factors considered important in the determination of the work-fertility interrelationship is the group of norms and beliefs governing family life. Particularly important are attitudes towards women as mothers and as workers. However, as far as fertility behaviour is concerned, employment of women in the service sector would be a more appropriate variable for consideration. The effect of participation of women in the service sector on fertility change would definitely be significant, both quantitatively and qualitatively. In this study, the major focus has been to capture the regional imbalances of women employment in the service sector contributing to the change in the childbearing process. But data on work participation rate of women at the district level are not available. Instead, data on percentage of total employment to the service sector (PWTSS) have been compiled from the 1991 Census and have been used as a rough indicator of district-level employment in the service sector since the imbalances among districts with respect to employment in the service sector of women would also be similar to that of PWTSS. Indirectly, PWTSS is assumed to be enough to capture the inequality of women's employment in the service sector among districts.

Per Capita Income

Of the many choices available for economic indicators affecting fertility, per capita income is one of the most widely accepted indicators. This captures people's capacity for daily consumption necessary for minimum calorie intake. Economic independence affects fertility in several ways. This can also be attributed in meeting the need for reproductive health care and family planning. But district-level per capita income data are nowhere compiled in the Indian Official System. In fact, no such attempt has been made in this direction. National sample surveys also do not have enough coverage for district-level income data. As a proxy for district-level per capita income data, per capita bank deposits (PCBD) have been used. It may be argued that bank deposits, in general, reflect the urban characteristics of the population. Since 1969, when the major banks in India were nationalised, the banking system has grown tremendously in terms of geographical reach and functional spread. It plays a crucial role in mobilising savings and capital accumulation through institutional savings. Each district of the major states is well connected by a banking network and the majority of the people have access to banking facilities. Per capita bank deposits indirectly capture the potential income capacity that is with the banking system. The disparity in individual income among districts should therefore be reflected in their power to capture savings also. As far as district imbalances are concerned, with respect to per capita income, PCBD may well serve the purpose. These data are taken from the CMIE publication on district profile.

Age at Marriage

The date of entry into first union is an important milestone in a woman's life; it represents not only a major change in the composition of her family but also the beginning of regular exposure to the risk of childbearing. Although marital unions form the essential conditions for childbearing and child rearing throughout the world, the structure, associated norms and customs, as well as the initial timing, prevalence and stability of unions vary widely. A trend towards delay of first marriage has been illustrated in developing countries with widely different economic, social and cultural configurations (Smith 1984; United Nations 1987; NFHS 1992–93). The increase in the age at entry into unions has been credited with a large share of the observed fertility decline. Broadly speaking, social and economic changes such as increased schooling for women and likely to be more urbanised, lead both to delays in marriage and to decline in marital fertility. These characteristics are associated with higher contraceptive use; thus, populations with later ages at marriage may show fairly low levels of fertility, not only because of their lost reproductive years but also because of deliberate limitation of marital fertility. Thus, indirectly, delayed age at marriage may enhance the motivation for family planning after marriage. This complicates the interpretation of relationships between the timing of entry into regular sexual exposure and completed fertility, particularly when such findings are to be adapted for policy application. Nevertheless, age at first birth can be viewed as a proximate determinant of fertility. More specifically, this may be viewed as a sufficient condition for fertility decline but may not be regarded as a necessary condition. There are large inter-state variations in age at marriage. The data on district-level age at marriage (MARGE) for 1991 have been taken from the 1991 census.

Some important points are worth noting. We have excluded those districts that are found to be outliers, *viz*., Greater Mumbai, Calcutta, Chennai, Hyderabad, etc. because of the highly influential characteristics of selected variables. In particular, these districts had 100 per cent urbanisation. Inclusion of these districts may distort the true picture of the states and may affect the results of the discriminant analysis.

The average and the standard deviation of these variables are presented in Statements 1 and 2. The average values of these variables have been computed on the basis of district-level observations and after excluding outlier districts, wherever applicable, and thus should not exactly tally with the values reported for a state. For example, the female literacy of Rajasthan, at 18.76 per cent in 1991 (as used here), is the average female literacy of all districts in Rajasthan, which is different from the data reported at the state level at 20.8 per cent. Due to non-availability of district-wise authentic data, no indicator was used on the provision of health and family welfare. However, the importance of health and family welfare services on population transition need not be overemphasised. In the discriminant analysis, we have assumed common population variance-covariance matrix for all the groups, and the sample pooled variance-covariance matrix is taken as the estimate.

Empirical Analysis

States are classified on the basis of TFR, IMR and NICR [natural rate of increase = (CBR-CDR)]. State-wise values of these indicators are presented in Table 1.

Table 1: State-wise Classification Indicators – 1992

State	TFR	IMR	NICR@
Kerala	1.7	17	11.4
Tamil Nadu	2.2	66	12.3
Andhra Pradesh	2.9	79	17.3
Karnataka	2.9	82	17.8
Maharashtra	2.9	67	17.4
West Bengal	2.9	71	16.4
Himachal Pradesh	3.1	69	19.3
Orissa	3.1	118	16.1
Punjab	3.1	61	18.9
Gujarat	3.2	72	18.9
Assam	3.4	83	20.4
Haryana	3.8	79	23.3
Madhya Pradesh	4.4	109	22.0
Rajasthan	4.5	94	24.4
Bihar	4.6	84	21.4
Uttar Pradesh	5.2	102	23.5

Note: @ NICR = CBR – CDR. *Source*: Sample Registration System 1993

Table 2 summarises the classification status of the states. There are four groups; group I consists of two states, *viz.*, Kerala and Tamil Nadu comprising 34 districts; group II consists of 7 states involving 130 districts; group III with 52 districts from 3 states; and group IV having 177 districts from 4 states. It can be seen that cluster analysis of states itself imposes some grouping pattern, which clearly depicts different stages of transition among groups. While group I states have almost completed fertility transition, group IV states are yet to make a significant dent in fertility transition.

Table 2: Classification of States by Average Linkage Clustering Method

Group	States	Range of variables					
		TFR		IMR		NICR	
		Min.	Max.	Min.	Max.	Min.	Max.
I	Kerala, Tamil Nadu	1.7	2.2	17	66	11.4	12.3
II	Gujarat, Andhra Pradesh, Maharashtra, West Bengal, Himachal Pradesh, Karnataka, Punjab	2.9	3.2	61	82	16.4	18.9
III	Orissa, Haryana, Assam	3.1	3.8	79	118	16.1	23.3
IV	Bihar, Madhya Pradesh, Rajasthan, Uttar Pradesh	4.4	5.2	84	109	21.4	24.4

The results of the discriminant analysis of 393 districts revealed four distinct classes. It may be mentioned that discriminant analysis may not yield the same number of discriminant functions as the number of groups considered earlier. Here we have found four discriminant functions specifying each group, which are tabulated in Table 3.

Table 3: Group-Wise Coefficients of Discriminant Functions

Variables	Groups			
	I	II	III	IV
CONSTANT	-207.46690	-166.39021	-151.61385	-146.34516
RDEV	-0.05215	-0.05163	-0.03270	-0.07346
URBAN	0.39229	0.37287	0.19823	0.34497
FLIT	0.50469	0.28830	0.20609	0.08897
PWTSS	-0.27016	-0.22231	0.20695	-0.02784
PCBD	-0.00626	-0.00517	-0.00578	-0.00466
MARGE	21.20945	19.44956	18.35063	18.53573
No. of districts	34	130	52	177
Prior Probabilities	0.086514	0.330789	0.132316	0.450382

The group-wise discriminant functions clearly reveal the impact of socio-economic variables to determine the stages of fertility transition for each group. The coefficient of each variable in the discriminant function indicates its relative weight or importance among all variables determining the discriminant function. On the other hand, prior probabilities put corresponding loadings of each group. As Table 3 shows, group IV had the highest prior probability followed by group II, as determined by the corresponding number of districts.

Female literacy is found to be the second most dominant factor to characterise the fertility transition in India after age at marriage. Urbanisation is the third most important variable in the selection and specification process. These three variables have consistently positive loadings on the discriminant functions. All other variables turn out to be not as significant as female literacy and age at marriage for the discrimination purposes. It is interesting to note that fertility behaviour of India is explained more by social characteristics than by corresponding economic counterparts. The coefficients of female literacy and age at marriage of group I states were 0.505 and 21.21 respectively. The most important feature of the results is that the impact of these two variables gets reduced as one moves from group I to group IV, i.e. from strong transition states to weak transition states. In other words, the states in group I have achieved a level of female literacy and age at marriage that could hasten the fertility transition process. The states in other groups, group IV in particular, have to achieve a bare minimum of female literacy and age at marriage to control the unprecedented growth of population. The existing level of these two may not be good enough to slow down the tempo of fertility. One needs to remember that female literacy and age at marriage are two highly positively correlated variables. Indeed, one may generally expect literacy to result in higher age at marriage. Alternatively speaking, literacy is a necessary condition for raising age at marriage and may not be a sufficient condition as the stringent legal framework may lead to a rise in age at marriage without recourse to the literacy path. Therefore, to influence the two crucial social variables, honest necessary efforts are to be made quickly as a part of the national movement.

The districts classified in each group are not likely to have exact similarities. That is, all districts in a particular group may not have the same characteristics to belong to the same group. More clearly, there can be some misclassified districts in each group, which are to be re-substituted and redistributed after adopting the discriminant procedure. For example, some districts in Tamil Nadu may not be of the same quality of Kerala districts. Therefore, such misclassified districts from group I should belong down the order in either group II, III or IV. If the initial group classification becomes satisfactory, it is expected to have less misclassification units, i.e., the greater the homogeneity of the class, the less are the chances of misclassification. A re-substitution summary of the districts is presented in Table 4.

It is observed that about 68 per cent of group I districts are correctly classified and the remaining 32 per cent of the districts should actually belong to group II. Thus, districts that are misclassified in group I are yet to register a significant fertility transition as compared with other group I districts. It is interesting to note that not a single district from group I was classified beyond group II districts. About sixty-five per cent of group II districts (84 districts out of 130) are correctly classified. Only 5 (about 4 per cent) districts in group II got berths in group I, while 41 districts (about 31 per cent) slipped to group III or IV, the majority (32 districts) of which should belong to group IV.

Table 4: Re-substitution Summary of Districts Using Linear Discriminant Function

From Group	Classified to Group				
	I	II	III	IV	Total
I	23	11	0	0	34
	(67.67)	(32.25)	(0.00)	(0.00)	(100.00)
II	5	84	9	32	130
	(3.85)	(64.62)	(6.92)	(24.62)	(100.00)
III	0	10	30	12	52
	(0.00)	(19.23)	(57.69)	(23.08)	(100.00)
IV	0	12	2	163	177
	(0.00)	(6.78)	(1.13)	(92.09)	(100.00)
Total	28	117	41	207	393
	(7.12)	(29.77)	(10.43)	(52.67)	(100.00)
Priors	0.0865	0.3308	0.1323	0.4504	1.000

Note: Figures in brackets indicate percentage share in total.

Among group III districts, about 58 per cent of the districts are correctly classified, while 42 per cent are misclassified either into group II or group IV. Of these, 19 per cent of the districts are of the level of group II districts. Classification summary for group IV districts is more accurate. Out of 177 group IV districts, 163 districts (92 per cent) are correctly classified. Only 14 districts are classified as over group IV, of which 12 belong to group II. The salient feature of group III and group IV districts is that not a single district acquired the status of belonging to group I. It may be mentioned that outliers that had exceptionally high values of selected variables are excluded from the purview of the study.

The following findings emerge from the results of the discriminant analysis: (1) The outcomes of this discriminant procedure are likely to be robust and reliable because of the very small error count in each group. (2) Out of 393 districts in India, only 28 districts (about 7 per cent) have achieved expected fertility transition. In other words, 93 per cent of Indian districts are still away from these 28 districts as far as fertility transition is concerned. (3) More than half of the Indian districts (52 per cent) belong to a very low category of fertility transition phase. Does it imply that the tempo of fertility in India has been arrested? The tempo has slowed down a bit, but the potential for higher fertility can never be ignored, especially for re-substituted group IV districts which account for over 65 per cent of India's population. The average total fertility rate of these districts is above 4. Therefore, on the average, fertility transition in India has not turned out to be a global phenomenon. If we look at the state-wise distribution matrix (Table 5) of the number of districts according to the level of fertility transition, as revealed by each group, a clear picture emerges.

Table 5: Group/State-wise Distribution Matrix of Districts After Classification

Group/state		Group I	Group II	Group III	Group IV	Total
Group I	Kerala	14	0	0	0	14
	Tamil Nadu	9	11	0	0	20
Group II	Gujarat	2	15	0	2	19
	Andhra Pradesh	0	6	1	15	22
	Maharashtra	1	23	0	5	29
	West Bengal	0	6	4	6	16
	Himachal Pradesh	0	8	3	1	12
	Karnataka	1	16	0	3	20
	Punjab	1	10	1	0	12
Group III	Assam	0	2	17	4	23
	Haryana	0	2	12	2	16
	Orissa	0	6	1	6	13
Group IV	Bihar	0	1	0	41	42
	Madhya Pradesh	0	9	0	36	45
	Rajasthan	0	0	0	27	27
	Uttar Pradesh	0	2	2	59	63

Kerala is the only state where the expected fertility transition has been already achieved. Tamil Nadu has also progressed towards the half-way mark since about half of the districts of Tamil Nadu effectively belong to group II. Next in the line are Punjab, Gujarat, Himachal Pradesh and Karnataka. The majority of the districts of Andhra Pradesh are in group II and are apparently similar to group IV districts. In fact, Andhra Pradesh is solely responsible for the misclassified districts in group II. Thus, the performance of fertility transition in Andhra Pradesh should be on a par with the group IV states. Orissa, in group III, has some peculiar characteristics. This state has registered some mixed mobility in the process of fertility transition; about half of the districts are above group III while the rest are below group III. The performance of group IV districts is really grim and awful. The condition of Rajasthan was most tragic. Among them, Madhya Pradesh had registered some improvement. All of these states are marked with very low female literacy and low age at marriage, besides being economically backward. Some exploratory analysis of these states with special reference to Rajasthan, Bihar and Uttar Pradesh is introduced below.

District-level re-substitution results of the misclassified districts after discriminant analysis are presented in the Appendix. The relative chance of each misclassified district belonging to each group in terms of their posterior probabilities is also presented in the Appendix. These posterior probabilities may serve as a very important indicator of the existing status of the fertility transition of a particular district. These are also extremely useful to predict the pace and mobility of fertility transition for each state. For example, Tiruvannamalai district, of Tamil Nadu, has a significantly high chance of transiting to group I, although effectively it belongs to group II. Similarly, if we see the misclassified group IV districts of Maharashtra we observe that most of them have a sizeable chance of shifting to at least group II. The same is the case of West Bengal. These indirectly imply that these states are heading rapidly towards fertility transition. In contrast to this, very few group IV districts are misclassified, i.e., there are few districts in group IV that are observed to have upward mobility. One needs to remember that any misclassification in group IV is nothing but an automatic upward shift of that particular district heading for better fertility transition.

A Lower Bound of Female Literacy and Fertility Transition

Increased education is among the aims of development planning in all countries, and policies that call for the increased integration of women into the development process specifically require that more educational opportunities be provided for women. The extent to which education affects not only fertility levels but other factors too that may be targets of development policy, such as maternal and child health, breast-feeding, contraceptive use, familial relationships and labour force participation, can have important implications for the achievement of population policy goals. Again the impact of education on fertility transition need not be overemphasised. Thus, we can safely assume that the fertility behaviour of women in India can also be sufficiently captured through the movement of female literacy. In other words, fertility transition must take place once a minimum level of female literacy is achieved, even if the other factors associated with fertility transition do not change. But improvement in fertility automatically derives from improvement in several socio-economic variables affecting fertility. The expected fertility transition may occur even before reaching the lower bound of female literacy. Therefore, this lower bound is nothing but the infimum literacy rate, i.e., maximum of the lower bounds of female literacy.

Here, we have tried to determine a lower bound of female literacy level by keeping other variables associated in the discriminant function at the state level. As it were, increase in women education is sufficient for fertility transition even if other factors do not improve. Therefore, as mentioned earlier, the estimated threshold value may signal some degree of overestimation. Fertility transition will thus occur before the threshold value is achieved.

The procedure is simple and straightforward. Let ‘x’ be the unknown threshold value of the female literacy which is to be estimated. For a state, let the variables used for discrimination be denoted by a vector Z= (RDEV, URBAN, x, PWTSS, PCBD, MARGE) where all variables, except ‘x’ are known. This vector will be allocated to group I, i.e., the state generating Z will shift to higher and expected fertility transition group if $d_i(Z)$ = largest of ($d_1(Z)$, $d_2(Z)$, $d_3(Z)$, $d_4(Z)$), where $d_i(Z)$ is the discriminant score of the i-th group, i=1, 2, 3 and 4. This has been explained in detail in the methodology section. The average values of all the variables are tabulated in Statement 1. State-wise disparities, as measured by the standard deviation and coefficient of variation of these variables, can be seen in Statement 2.

We have mentioned earlier that fertility transition of three states, viz., Rajasthan, Uttar Pradesh and Bihar belonging to group IV has been marginal and the pace of change is also very slow. These three states received special attention in this study and the computation of the lower bound of female literacy has been confined to these states only. The lower bound of female literacy of these states has been worked out as follows:

<u>State= Rajasthan, Lower bound = y, Group = IV, Target = Group I</u>

$d_1(y) > d_2(y)$, $d_1(y) > d_3(y)$ and $d_1(y) > d_4(y)$ imply that

$116.59 + 0.50469y > 132.04 + 0.28830y \Rightarrow 0.21639y > 15.45 \Rightarrow y > 71.3988$

$116.59 + 0.50469y > 133.86 + 0.20609y \Rightarrow 0.29860y > 17.27 \Rightarrow y > 57.8366$

$116.59 + 0.50469y > 139.92 + 0.08897y \Rightarrow 0.41572y > 23.33 \Rightarrow y > 56.1195$

⇨ $y > 71.3988$.

⇨ Hence, Rajasthan requires at least 71.39 per cent of female literacy to reach group I.

<u>State= Uttar Pradesh, Lower bound = x, Group = IV, Target = Group I</u>

$d_1(x) > d_2(x)$, $d_1(x) > d_3(x)$ and $d_1(x) > d_4(x)$ imply that

$124.15 + 0.50469x > 139.26 + 0.28830x \Rightarrow 0.21639x > 15.11 \Rightarrow x > 69.8276$

$124.15 + 0.50469x > 140.92 + 0.20609x \Rightarrow 0.29860x > 16.77 \Rightarrow x > 56.1621$

$124.15 + 0.50469x > 146.92 + 0.08897x \Rightarrow 0.41572x > 22.77 \Rightarrow x > 54.7724$

⇨ $x > 69.8276$.

⇨ Therefore, Uttar Pradesh requires at least 69.83 per cent of female literacy to reach group I.

State= Bihar, Lower bound = z, Group = IV, Target = Group I

$d_1(z) > d_2(z)$, $d_1(z) > d_3(z)$ and $d_1(z) > d_4(z)$ imply that

$128.88 + 0.50469z > 142.47 + 0.28830z \Rightarrow 0.21639z > 13.59 \Rightarrow z > 62.8073$

$128.88 + 0.50469z > 145.35 + 0.20609z \Rightarrow 0.29860z > 16.48 \Rightarrow z > 55.1898$

$128.88 + 0.50469z > 149.82 + 0.08897z \Rightarrow 0.41572z > 20.94 \Rightarrow z > 50.3772$

⇨ $z > 62.8073$.

⇨ Thus, Bihar needs at least 62.81 per cent female literacy to reach group I.

As expected, Rajasthan has a relatively greater requirement in female literacy than Uttar Pradesh and Bihar. The lesser requirement for Bihar should not be looked upon as a different phenomenon. Bihar had comparatively less values of RDEV, PWTSS and PCBD, which had negative loadings (probably insignificant) in the discriminant functions, which resulted in a lower numerator in the calculation and hence relatively lower value of the threshold value. And since the values of RDEV, PWTSS and PCBD are significantly low for Bihar, this will be reflected in the movement of female literacy growth. Therefore, the time requirement for Bihar to reach group I level may not yield something different as compared with Rajasthan and Uttar Pradesh. A time frame for the fertility transition of these states is devised in the next subsections. One point is again worth noting. The time for transition, as derived here, basically dictates the expected transition path if the current level of development prevails.

Time Required for Fertility Transition

Here the fertility transition between groups has been envisaged by modelling the female literacy movements. Suppose that L(t), the percentage of female literacy at time 't' follows a logistic law, i.e., $L(t) = K/[1+\exp.(a+bt)]$, where K = highest value of L(t) = 100.

The parameters 'a' and 'b' are estimated by the usual Pearl's method of estimating the logistic curve. Table 6 presents the relative movement of female literacy of the three selected states along with their desired target level. As pointed out earlier, the relative growth in female literacy of Bihar is comparatively less than that of Uttar Pradesh.

Table 6: Trend in Female Literacy Rate of Selected States

State	1961	1971	1981	1991	Target
Uttar Pradesh	8.43	12.46	16.33	25.31	69.828
Rajasthan	7.00	10.00	11.00	18.76	71.399
Bihar	8.00	10.00	14.00	22.50	62.807

Once the parameters 'a' and 'b' are estimated from the past data of L(t), the required time can be estimated from the following formula: t_0 = [log{K/L(target) - 1} – a_0]/b_0, where a_0 and b_0 are the respective estimates of a and b.

Table 7 presents state-wise parameters estimates and the time required for the transition of states belonging to lower groups to higher groups using the above formula.

Table 7: Time Requirement for Fertility Transition of Selected Group IV States to Group I

State	Parameter Estimate a	b	Target	Required time (yrs.)
Rajasthan	2.0907	-0.0625	71.399	48.08
Uttar Pradesh	1.6339	-0.0552	69.828	44.83
Bihar	1.8153	-0.0579	62.807	40.43

We observe that Rajasthan will take 48 years more to achieve the group I status, while Uttar Pradesh and Bihar will take about 45 and 40 years respectively to reach the group I stage of fertility transition. In other words, the expected fertility transition of India's most populated states may occur only after 2030 AD if the present rate of change of female literacy and related efforts prevails. Thus the results indicate that the faster the growth of female education, the quicker is the process of fertility transition.

The results clearly show that economic reforms alone are not sufficient unless accompanied by social commitments. Faster development requires government action to improve elementary education, and health care, and to remove barriers against certain sections of society, particularly women. Kerala has set an example and has shown the way. Its success has very little to do with economic growth because in spite of its social progress it has a sluggish economy and a very high level of unemployment. However, the importance of economic growth should not be understated.

Threshold of Female Literacy and Fertility Transition

Following the same logic as given by the United Nations (1963) for postulating the 'threshold hypothesis' regarding fertility transition and development, we can define 'threshold hypothesis' for fertility transition and female literacy. Fertility transition of a society is possible only when literacy of women surpasses the threshold; once that level is achieved, fertility is likely to decline faster and continue downward until it is stabilised. The proper derivation of the threshold depends on how best one can model the relationship between fertility and women's literacy level.

The Model

Let the total fertility rate (TFR) be denoted by f(x) corresponding to female literacy level 'x'. Graph 1 exhibits the plot of f(x) against log(x) where TFR and literacy level correspond to each major state in India for the year 1991. From the plot it is evident that as log(x) increases (i.e. 'x' increases), f(x) follows a gradual slow and flat exponential decay path which has some similarity to log-normal distribution. Therefore, we assume,

$f(x) = a*\exp\{c*(\log x - b)^2\}$, ...(2)

where a, b, c are constants, with $a > 0$, $b > 0$ and $c < 0$.

But our interest is to find out how far the total fertility rate changes with the change in women's literacy. By differentiating f(x) with respect to 'x', we get,

$g(x) = f'(x) = 2*a*c*\{(\log x - b)/x\}*\exp\{c*(\log x-b)^2\} = 2*c*\{(\log x - b)/x\}*f(x)$

Therefore, the change in TFR for the change in women literacy is given by

$g(x) = 2*c*\{(\log x - b)/x\}*f(x)$..(3)

Now, the threshold of women literacy is that value of 'x' for which g(x) is maximum. In other words, it is that value of 'x' for which $g'(x) = 0$ and $g''(x) < 0$.

Differentiating (3) with respect to 'x' and equating it to 0, we get

$f(x)\,[\{2*c*(\log x-b)/x\}^2 + (2*c)/x^2 + 2*c*(\log x-b)*(-1/x^2)] = 0$

After simplification, we have $2*c*(\log x-b)2 - (\log x-b) + 1 = 0$

$\Rightarrow 2*c*z2 - z + 1 = 0$, where $\log x-b = z$ (say)

$\Rightarrow z = \{1 \pm \sqrt{(1-8c)}\}/(4*c) \Rightarrow \log x - b = \{1 \pm \sqrt{(1-8c)}\}/(4*c)$

$\Rightarrow x = \exp[\{1 \pm \sqrt{(1-8c)}\}/(4*c)\} + b] = x_0$ (say)..(4)

Now for the second order condition we need to have $g''(x_0) < 0$.

Estimation of the Parameters a, b and c

As the function f(x) is non-linear in 'x', the linearity assumption of the classical regression analysis is no longer valid. So the normal equations by minimizing the residual sum of squares are of non-linear nature. Here we have estimated the parameters by the method of non-linear least square (NLINLS) iterative procedure, which is also known as Gauss-Newton method (the general algorithm for solving a system of non-linear equations was developed by Gauss and Newton). Therefore, estimation of the parameters by the iterative procedure may not be unbiased and minimum variance estimators (UMVE). Accuracy of the parameters is to be judged by either plotting the original data with the predicted values or by seeing their asymptotic standard errors.

Table 8 reports the non-linear least square estimates of the parameters, while graph 2 shows the plot of observed and predicted values of TFR against female literacy.

Table 8: Non-Linear LS Estimate of the Parameters

Parameter	Estimate	Asymptotic Standard Error	Asymptotic 95% Confidence Interval Lower	Upper
A	5.10503	1.8869	1.0285	9.1815
B	3.02787	1.3271	-0.4884	5.2457
C	-0.25789	0.3016	-0.9096	0.3938

Note: Parameters are estimated by using SAS Software (Version 6.12)

Graph 1: Plot of TFR against States Female Literacy Rates (Literacy is in Logarithmic Term)

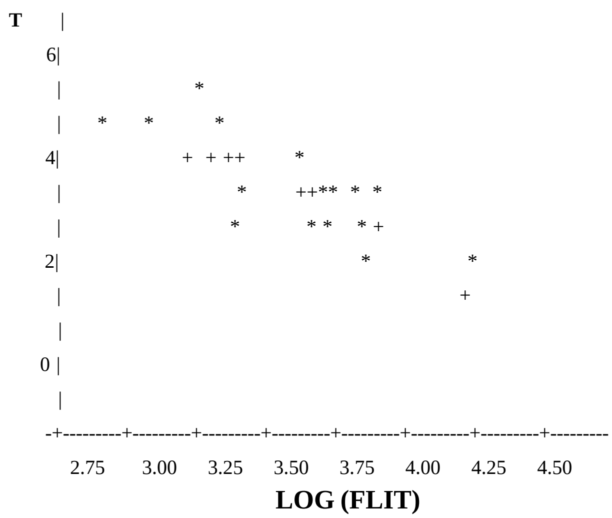
```
     6|
      |
      |                 A
      |   A     A          A
     4|                              A
      |                                   A
 TFR  |                       A A         AAA  B   A
      |
     2|                                            A
      |                                                                    A
      |
      |
     0|
     -+---------+---------+---------+---------+---------+---------+---------+---------
         2.75      3.00      3.25      3.50      3.75      4.00      4.25      4.50
```

LOG(FLIT)

Note: Legend: A = 1 observation, B = 2 observations. *Source*: 1991 Census.

It is seen that asymptotic standard error of the parameters is quite low and the model gives a good fit with the observed data. Therefore, the estimates of the parameters a, b and c are very reliable and all have the expected sign. Putting the estimated value of a, b and c in equation (3), we get x = 1.44412 (corresponding z = -2.666037) and x = 42.73819 (corresponding z = 0.727226). Two values of x identify two optimum values of ‘x’, i.e., g(x) attains a maximum at x = 42.74 and a minimum at x = 1.44. It can be shown that $g^{//}(42.74) < 0$. Therefore, the rate of decline of TFR with respect to female literacy would be the maximum when the female literacy rate of India would reach the level of 42.74 per cent. Once this level is achieved fertility will decline faster towards the stability of the population growth. Hence, the estimated threshold of female literacy with respect to fertility transition in India is found to be 42.7 per cent.

Graph 2: Plot of Observed TFR *vis-a-vis* Plot of Predicted TFR Against States Female Literacy Rates



Notes: 1. 9 observations are hidden because of their very close prediction.
2. For original TFR '*' is used and for predicted TFR '+' symbol is used.

States are arranged according to the level of female literacy and are presented in Table 9. It can be seen from the table that six states have already surpassed the threshold of female literacy and these states have actually registered faster fertility decline in the recent period. Another three states are in the vicinity of the threshold.

Table 9: Distribution of States According to Female Literacy

Range of female literacy (per cent)					
< 20	20-30	30-40	40-43	43-60	60+
Rajasthan	Bihar	Orissa	Assam,	Karnataka,	Kerala
	Uttar Pradesh	Andhra Pradesh	Himachal Pradesh	Gujarat	
	Madhya Pradesh	Haryana	West Bengal	Maharashtra	
				Punjab	
				Tamil Nadu	

Source: Census of India, 1991

But states like Rajasthan, Bihar, Uttar Pradesh, and Madhya Pradesh have to search for a strategic plan to raise the female literacy level to at least 43 per cent to trigger off a faster change in the level of fertility.

Summary and Conclusions

One of the main objectives of this paper was to classify major Indian states into some homogeneous groups characterising different levels of fertility transition. The observed differential in fertility between different groups, as determined by the cluster-cum-discriminant analysis, by taking the district as a unit of observation, indicate the urgent need for social uplift of women, especially in respect of female education and age at marriage. The economic variables, on the other hand, are found to be less important for the existing fertility differential between states. The comparative overview given here may clarify the reasons for the typical relationship between fertility and other socio-economic variables to determine which aspects of a district's population dynamics are to be addressed first. The findings, in general, strongly confirm the need for district-level planning by focusing more on the specific need and urgency of the problems associated with the population change.

It is important to note that the emphasis here has been on the relationships between the levels of female education and fertility transition. Two separate approaches are followed here: (a) The first approach determines the maximum of a lower bound (infimum) of female literacy required for below replacement fertility using Bayesian discriminating procedure, while (b) a threshold of female literacy is derived by modelling the relationship between female literacy and total fertility rate in the second approach. Both the approaches have tried to quantify the magnitude of female literacy required for fertility transition in India. The findings suggest that the threshold of female literacy for a faster decline in fertility in India is about 43 per cent; once that level is achieved the fertility rate will register a faster decline towards the stability of the population.

The evidence presented in this article regarding the association between education, fertility and fertility-related variables are, in general, consistent with the general perceptions. Apart from data limitations at the district level, the results reported in this paper depend heavily on the validity of the models and the variables used in the analysis to capture fertility differential among regions. However, some crucial points are worth mentioning. First, the role of family planning programme in declining fertility has not been incorporated in the district-level analysis. It is well known that the family planning programme in India is by and large successful in reducing the span of demographic transition in many states. There are factors that directly or indirectly influence the adoption of family planning methods. Secondly, during the nineties, the female literacy rate and other socio-economic indicators have markedly improved. The analysis based on 1991 data in determining the time required for fertility transition would be on the higher side. The findings of this study, thus, are limited in that sense. However, further research on the extent to which education of women affects not only fertility levels but other factors that may be the target of development policy, such as maternal and child health, and contraceptive use, would have important implications for the achievement of population policy goals.

Statement 1: State-wise Distributions of the Average of the Selected Indicators

Group/State	RDEVIN	URBAN	FEMLIT	PWTSS	PCBD	MARGE
Group I						
Kerala	113.57	23.37	85.55	32.14	3885.36	19.11
Tamil Nadu	109.9	28.79	50.45	19.64	1894.65	18.4
Group II						
Andhra Pradesh	86.64	22.75	30.08	16.40	1490.36	15.87
Gujarat	106.12	31.37	47.49	22.05	3507.06	18.32
Himachal Pradesh	73.75	10.82	42.31	20.67	2933.17	17.43
Karnataka	111.85	26.10	43.57	18.06	2388.00	17.04
Maharashtra	78.79	24.99	47.23	16.60	1483.07	16.46
Punjab	200.17	27.91	50.35	29.16	5715.17	18.82
West Bengal	68.31	20.32	41.22	21.71	1277.38	16.33
Group III						
Assam	62.70	10.21	43.53	19.64	1039.04	15.07
Haryana	149.38	23.80	40.75	28.24	2981.31	16.74
Orissa	64.23	12.96	31.30	15.65	1004.85	17.24
Group IV						
Bihar	43.12	12.56	22.50	16.32	1184.38	15.94
Madhya Pradesh	70.57	22.19	28.26	14.55	1365.78	15.49
Rajasthan	63.33	20.70	18.76	17.57	1388.44	15.69
Uttar Pradesh	70.90	18.59	27.14	18.30	1948.87	16.27

Statement 2: State-wise Variations of Selected Indicators

States		RDEVIN	URBAN	FEMLIT	PWTSS	PCBD	MARGE
Kerala	STD	27.18	14.67	5.52	7.54	2500.29	0.74
	CV	23.94	62.78	6.45	23.47	64.35	3.89
Tamil Nadu	STD	31.57	12.46	9.91	4.49	969.44	0.71
	CV	28.72	43.26	19.65	22.85	51.17	3.87
Andhra	STD	17.41	8.67	8.59	4.25	535.76	1.00
Pradesh	CV	20.09	38.10	28.57	25.89	35.95	6.30
Gujarat	STD	46.72	16.82	12.83	8.35	2499.36	0.48
	CV	44.02	53.59	27.02	37.86	71.27	2.63
Himachal	STD	26.70	4.68	12.39	4.78	1594.18	1.44
Pradesh	CV	36.20	43.22	29.28	23.14	54.35	8.25
Karnataka	STD	52.24	14.72	12.55	7.55	2591.86	0.97
	CV	46.70	56.41	28.82	41.81	108.54	5.69
Maharash-	STD	23.09	13.73	10.86	5.87	1114.33	0.74
tra	CV	29.30	54.95	22.99	35.36	75.14	4.49
Punjab	STD	42.99	8.55	9.59	5.40	3654.67	0.25
	CV	21.48	30.62	19.05	18.53	63.95	1.33
West	STD	17.12	14.36	11.91	7.24	757.17	0.66
Bengal	CV	25.06	70.70	28.88	33.35	59.28	4.06
Assam	STD	26.43	6.84	7.70	6.58	959.54	0.49
	CV	42.15	66.95	17.69	33.50	92.35	3.25
Haryana	STD	56.37	8.94	7.77	6.39	1148.77	0.49
	CV	37.74	37.57	19.06	22.61	38.53	2.93
Orissa	STD	12.60	7.11	12.06	5.33	690.25	0.45
	CV	19.62	54.84	38.52	34.09	68.69	2.63
Bihar	STD	16.30	11.31	7.53	9.41	1059.07	0.57
	CV	37.80	90.03	33.47	57.70	89.42	3.61
Madhya	STD	27.04	15.10	10.69	8.50	1490.65	0.87
Pradesh	CV	38.33	68.04	37.84	58.45	109.14	5.63
Rajas-	STD	25.88	9.65	6.40	5.68	732.81	0.87
than	CV	40.86	46.62	34.12	32.32	52.78	5.57
Uttar	STD	28.60	14.91	11.68	9.40	1812.7	0.75
Pradesh	CV	40.34	80.22	43.06	51.36	93.01	4.59

Note: STD = Standard Deviation, CV = Coefficient of Variation

Appendix

Group/State/District-wise Resubstitution Results of the Misclassified Observations Using Linear Discriminant Function

Group/Districts	Original Group	Classified as		Posterior Probabilities I	II	III	IV
Group - I							
Kerala							
Tamil Nadu							
Coimbatore	1	2	*	0.1637	0.8349	0.0002	0.0012
Dharmapuri	1	2	*	0.0120	0.8032	0.0648	0.1200
Dindigulanna	1	2	*	0.0596	0.8702	0.0303	0.0399
Kamarajar	1	2	*	0.3876	0.6096	0.0007	0.0020
North Arcot	1	2	*	0.1790	0.7764	0.0320	0.0125
Pasumpan	1	2	*	0.2695	0.7179	0.0019	0.0107
Periyar	1	2	*	0.0877	0.8919	0.0052	0.0152
Salem	1	2	*	0.0313	0.9350	0.0092	0.0245
South Arcot	1	2	*	0.0579	0.8963	0.0167	0.0291
Tiruchirappalli	1	2	*	0.2769	0.7156	0.0029	0.0047
Tiruvannamalai	1	2	*	0.4093	0.5809	0.0080	0.0018
Group – II							
Gujarat							
Amreli	2	1	*	0.5605	0.4368	0.0014	0.0013
Banaskantha	2	4	*	0.0032	0.4259	0.0330	0.5380
Kachchh	2	4	*	0.0000	0.1918	0.0018	0.8064
Valsad	2	1	*	0.7702	0.2289	0.0007	0.0002
Andhra Pradesh							
Adilabad	2	4	*	0.0000	0.0427	0.0071	0.9502
Anantapur	2	4	*	0.0004	0.4774	0.0294	0.4927
East Godabari	2	3	*	0.0004	0.3581	0.3697	0.2718
Karimnagar	2	4	*	0.0000	0.0476	0.0418	0.9106
Khammam	2	4	*	0.0001	0.3889	0.0405	0.5705
Kurnool	2	4	*	0.0001	0.2931	0.0192	0.6876
Mahbubnagar	2	4	*	0.0000	0.0217	0.0175	0.9608
Medak	2	4	*	0.0000	0.0338	0.0457	0.9205
Nalgonda	2	4	*	0.0000	0.0594	0.1181	0.8225
Nizamabad	2	4	*	0.0000	0.0383	0.0613	0.9004
Prakasam	2	4	*	0.0000	0.2833	0.0471	0.6696
Rangareddy	2	4	*	0.0001	0.3078	0.0291	0.6631
Srikakulam	2	4	*	0.0001	0.1720	0.0522	0.7757

Vishakhapatnam	2	4	*	0.0000	0.1712	0.0170	0.8118
Vizianagaram	2	4	*	0.0002	0.1731	0.0565	0.7702
Warangal	2	4	*	0.0000	0.0618	0.0593	0.8789
Maharashtra							
Amravati	2	1	*	0.7207	0.2785	0.0003	0.0005
Bid	2	4	*	0.0002	0.3662	0.0487	0.5849
Gadchiroli	2	4	*	0.0001	0.3887	0.0752	0.5360
Jalna	2	4	*	0.0002	0.4030	0.0254	0.5714
Nanded	2	4	*	0.0001	0.2469	0.0312	0.7218
Parbhani	2	4	*	0.0001	0.3053	0.0332	0.6614
West Bengal							
Bankura	2	4	*	0.0004	0.3500	0.2894	0.3602
Birbhum	2	3	*	0.0005	0.3056	0.4216	0.2722
Jalpaiguri	2	4	*	0.0006	0.1708	0.2901	0.5386
Koochbihar	2	3	*	0.0001	0.1223	0.6024	0.2752
Maldah	2	4	*	0.0000	0.0722	0.1587	0.7692
Murshidabad	2	4	*	0.0001	0.1450	0.2525	0.6024
Nadia	2	3	*	0.0017	0.3491	0.4334	0.2158
Purulia	2	4	*	0.0000	0.0459	0.0715	0.8826
S24 Parganas	2	3	*	0.0018	0.1989	0.5981	0.2012
West Dinajpur	2	4	*	0.0002	0.2488	0.0742	0.6768
Himachal Pradesh							
Bilaspur	2	3	*	0.0072	0.4635	0.4890	0.0403
Chamba	2	4	*	0.0001	0.1535	0.0827	0.7637
Hamirpur	2	3	*	0.0064	0.4206	0.5277	0.0453
Kangra	2	3	*	0.0155	0.3403	0.6032	0.0411
Karnataka							
Bidar	2	4	*	0.0000	0.1576	0.1223	0.7201
Gulbarga	2	4	*	0.0000	0.1200	0.0241	0.8559
Kodagu	2	1	*	0.6721	0.3231	0.0047	0.0000
Raichur	2	4	*	0.0000	0.1497	0.0437	0.8065
Punjab							
Hosiarpur	2	1	*	0.7428	0.0439	0.2132	0.0001
Rupnagar	2	3	*	0.0022	0.0689	0.9246	0.0043
Group - III							
Assam							
Darrang	3	2	*	0.0040	0.5016	0.4554	0.0391
Dhubri	3	4	*	0.0000	0.0229	0.2144	0.7626
Karbianglong	3	4	*	0.0001	0.2703	0.1520	0.5776
Kokrajhar	3	4	*	0.0000	0.0783	0.2892	0.6324

Marigaon	3	2	*	0.0012	0.4245	0.3695	0.2047
Tinsukia	3	4	*	0.0001	0.2429	0.1171	0.6399
Haryana							
Faridabad	3	2	*	0.0016	0.7186	0.1200	0.1598
Kaithal	3	4	*	0.0001	0.2500	0.3421	0.4078
Sirsa	3	2	*	0.0014	0.5330	0.3954	0.0702
Yamunanagar	3	4	*	0.0010	0.3946	0.0830	0.5214
Orissa							
Balangir	3	4	*	0.0003	0.2982	0.0486	0.6529
Baleshwar	3	2	*	0.0072	0.5250	0.3351	0.1327
Dhenkanal	3	2	*	0.0085	0.6789	0.1350	0.1777
Ganjam	3	4	*	0.0003	0.2776	0.0705	0.6516
Kalahandi	3	4	*	0.0000	0.0680	0.0236	0.9083
Kendujhar	3	2	*	0.0132	0.7261	0.0313	0.2293
Koraput	3	4	*	0.0000	0.0873	0.0105	0.9022
Mayurbhanj	3	4	*	0.0003	0.2865	0.0536	0.6597
Phulbani	3	4	*	0.0003	0.2723	0.0435	0.6839
Puri	3	2	*	0.0089	0.4713	0.4235	0.0963
Sambalpur	3	2	*	0.0044	0.7263	0.0441	0.2252
Sundargarh	3	2	*	0.0189	0.8165	0.0098	0.1548
Group - IV							
Bihar							
Purbisinghbhum	4	2	*	0.0025	0.7694	0.0002	0.2279
Madhya Pradesh							
Balaghat	4	2	*	0.0043	0.8352	0.0333	0.1271
Betul	4	2	*	0.0040	0.7813	0.0057	0.2089
Chhindwara	4	2	*	0.0012	0.6811	0.0181	0.2996
Durg	4	2	*	0.0006	0.7551	0.0065	0.2378
Eastminar	4	2	*	0.0004	0.5567	0.0143	0.4287
Jabalpur	4	2	*	0.0004	0.5844	0.0323	0.3829
Narsimhapur	4	2	*	0.0011	0.6950	0.1807	0.1232
Raigarh	4	2	*	0.0008	0.4941	0.0188	0.4863
Seoni	4	2	*	0.0005	0.4871	0.0342	0.4783
Rajasthan	4						
Uttar Pradesh							
Garhwal	4	3	*	0.0009	0.1525	0.6272	0.2195
Kanpurdehat	4	2	*	0.0069	0.7412	0.1431	0.1088
Nainital	4	2	*	0.0021	0.6652	0.1563	0.1765
Pithoragarh	4	3	*	0.0001	0.1289	0.4969	0.3741

Note: Misclassified districts are indicated by *

References

Anderberg, M.R. (1973). *Cluster Analysis for Applications*. New York: Academic Press.

Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Methods* (2nd Edition). New York: John Wiley.

Bose, A. (1994). *Demographic Zones in India*. Delhi: B.R. Publishing Corporation.

Caldwell, John C. (1983). Direct Economic Costs and Benefits of Children. In Rodolfo A. Bulatao, and Ronald D. Lee (eds.), with Paula E. Hollerbach and John Bongaarts, *Determinants of Fertility in Development Countries, Vol.1, Supply and Demand for Children.*

Chatfield, C., and Collins, A.J. (1980). *Introduction to Multivariate Analysis*. London: Chapman and Hall.

CMIE (1994). *Districts Profiles*. Economic Intelligence Unit.

Easterlin, R.A. (1978). *The Economics and Sociology of Fertility: A Synthesis*. In Charles Tilly (ed.). *Historical Studies of Changing Fertility.* Princeton, New Jersey: Princeton University Press.

Guru, M. (1992). *Fertility Destabilisation in India: An Analytical Examination*, Unpublished Ph.D. Thesis. Mumbai: International Institute for Population Sciences.

Johnson R.A., and Wichern D.W. (1996). *Applied Multivariate Statistical Analysis*. New Delhi: Prentice Hall of India Pvt. Ltd.

Mitra, A. (1965). *Census of India*, 1961, Vol.1, Part I-A (1), Levels of Regional Development, New Delhi: Office of the Registrar General.

National Family Health Survey (MCH and Family Planning), 1992–93, India, International Institute for Population Sciences, Bombay, August 1995.

Navaneetham, K. (1993). Mortality Decline in India: An Analysis of Regional and Temporal Variations. *Demography India*, 22(1): 53–63.

Office of the Registrar General (1993). *Sample Registration System: Fertility and Mortality Indicators 1992*, New Delhi.

Oppong, C. (1983). Women's Roles, Opportunity Costs and Fertility. In Rodolfo A. Bulatao, and Ronald D. Lee (eds.), with Paula E. Hollerbach and John Bongaarts. *Determinants of Fertility in Developing Countries, Vol.1, Supply and Demand for Children.*

Pathak, P. (1991). *Diagnostic Regional Analysis on the Shortfalls in Development and Utilisation of Human Resources in India: A Statistical Study on the Spatial Variations by Districts in India*, Unpublished Ph.D. Thesis, Indian Statistical Institute, Calcutta.

Report of the National Research Council Committee on Population and Demography, Panel on Fertility Determinants. New York: Academic Press, pp. 458–93.

Sen, A. (1997). *Indian Development: Selected Regional Perspectives*, edited by Jean Dreze and Amartya Sen, Oxford University Press.

Sen, A. and Jean Dreze (1998). India: Economic Development and Social Opportunity. In Sen, A., and Jean Dreze (eds.). *The Amartya Sen & Jean Dreze Omnibus*.

Smith, P.C. (1984). The Impact of Age at Marriage and Proportions Marrying on Fertility. In Rodolfo A. Bulatao, and Ronald D. Lee, (eds.), with Paula E. Hollerbach and John Bongaarts, *Determinants of Fertility in Developing Countries, Vol.2, Fertility Regulation and Institutional Influences.*

U. Datta Roy Choudhury (1995). *National Income Accounting, Inter-Regional and Intra-Regional Variations*, Chapter 20, MacMillan India Limited, pp.369–87.

United Nations (1963). *Population Bulletin of the United Nations*, No.7. New York: United Nations.

——— (1987). *Fertility Behaviour in the Context of Development: Evidence from the World Fertility Survey.* Department of International Economic and Social Affairs. New York: United Nations.