STUDIES

No. 14-02

Demographic Regimes in Rural India

Aalok Ranjan Chaurasia

'Shyam' Institute

This page is intentionally left blank

DEMOGRAPHIC REGIMES IN RURAL INDIA

AN ANALYSIS OF 2011 CENSUS DATA

ABSTRACT

The paper uses primary census abstract to analyse demographic situation in villages of India. The analysis is based on six indicators: 1) ratio of the population 0-6 years to the population 7 years and above; 2) ratio of the population 0-6 years to females 7 years and above; proportion of Scheduled Castes; 4) proportion of Scheduled Tribes; 5) proportion of population aged 7 years and above literate; 6) proportion of workers to population aged 7 years and above; and 7) proportion of main workers to total workers. The seven indicators have been grouped into three population domains through the application of factor analysis technique. An application of cluster analysis using the factor scores in the three population domains suggest that 597619 villages of the country can be grouped into six clusters with clear geographical patterns and distinct demographic situation. The analysis based on the 2011 population census is reported here and the analysis based on 2001 and 1991 population census is currently in progress to analyse how the village level demographic situation changed over the 20 years between 1991 and 2011.

INTRODUCTION

India is a country of villages. According to the 2011 population census, there were more than 640 thousand villages of varying population size in the country. Population living in villages, known as the rural population, constituted approximately 69 per cent of India's population enumerated at the 2011 population census (Government of India, 2011). Villages in India form an integral part of industrial, commercial and agricultural landscape of the country since ancient times. People living in villages are conditioned by a very diverse and heterogeneous social, religious, cultural, natural and economic environment that shapes their capabilities and capacities and their participation in the social and economic production system. The village social and economic production system, in turn, is an important component of the social and economic production system of the country.

Despite the importance of village society and economy in the national society and economy, little is currently known about distinguishing features of the village population in India. An understanding of village population stock and activities of the village people is necessary to understand the role villages play in the demographic dynamics and development processes of the country. It is however not easy to determine distinguishing characteristics of village population in India. Evidence on the demographic characteristics of villages in India is limited. The only source of village level demographic data in India is the Primary Census Abstract (PCA) prepared on the basis of decennial population census. These data, however, have rarely been used to analyse village demographic situation. One reason is that PCA does not facilitate estimation of conventional measures of demographic situation such as level of fertility, mortality and migration. Still, an understanding of village demography can be made through some innovative analysis of these data.

This paper uses data available through PCA of the 2011 population census to characterise population residing in Indian villages. The purpose is to identify main demographic regimes that prevail in the villages of India. The context of demographic regime is different from the context of demographic transition that was first propounded 70 years ago (Notestein, 1945; Davis 1945) and the search for an understanding of demographic transition in different social and economic settings has been the central preoccupation of modern demography (Demeny, 1972). The concept of demographic regime was introduced by Landry (1934). It refers to a particular combination of interrelated demographic characteristics that

pertains in a given population (Pressat, 1985). It is argued that patterning of demographic regime at the village level has implications for grass roots level population and development planning and programming. The characteristics of the village population may be seen as being shaped by the village social and economic production system. The key question that the present paper attempts to answer is how the demographic situation varies across the villages of the country and what are the social and economic implications of the prevailing demographic situation at the village level.

The analysis is based on the primary census abstract (PCA) of each village of the country which presents demographic, social economic and cultural profile of each village. PCA comprises number of households, enumerated population, population in the age-group 0-6 years, total number of literates and illiterates, total workers, main and marginal workers classified by four broad industrial categories - cultivators, agricultural labourers, household industry workers, other workers - and non workers separately for males and females. Data available through PCA permits estimation of selected indicators that reflect the demographic situation in the village.

THE CONCEPT OF VILLAGE IN INDIAN POPULATION CENSUS

The concept of village in the Indian population census is essentially a residual concept. Indian population census classifies the enumerated population into rural population and urban population. Urban population is defined as population living in towns or urban habitations whereas rural population is population living in villages or rural habitations. A town or an urban habitation has been well defined in the population census but a village has never been defined in the same manner. All habitations which have not been classified as a town or an urban habitation are classified as rural habitations. The basic unit of enumeration in the rural areas is a village which follows limits of a revenue village which has definite surveyed boundaries (Government of India, 1951). A revenues village, however, may or may not have any habitation. At the same time, a revenue village may have more than one habitations. The choice of revenue village as the basic unit for population enumeration is driven more by administrative convenience than by settlement logic.

At the 2011 population census, there were 640949 villages of varying population size in India out of which 43334 (6.8 per cent) villages were

uninhabited. In the remaining 597615 (93.2 per cent) villages, the enumerated population ranged from just one person to 66062 persons. Figure 1 gives the distribution of villages by the size of the population. Around 14 per cent of the villages in the country had a population less or equal to 200 inhabitants whereas about 4 per cent had a population of more than 5000 at the 2011 population census. Nearly half of the inhabited villages had a population ranging from 501 to 2000 while another about one fifth villages had a population ranging from 201 to 500. Around one fifth of the villages had a population of more than 2000. There were 453 villages in the country which had a population of only one inhabitant at the time of 2011 population census. In addition, in 4145 villages, the population enumerated was less than 10. On the other hand, there were five villages which had a population of more than 50 thousand with village Fursungi in district Pune of Maharashtra being the largest village of the country with a population of 66062. Village Edakkara in district Mallapuram of Kerala also had a population of more than 60000 at the 2011 population census. Moreover, population of 28 villages ranged between 40-50 thousand. In all, there were 23317 (3.6 per cent) villages in the country where the population was enumerated to be more than 5000 at the 2011 population census. These habitations, despite having very large population, could not be classified as a town as they could not meet the criteria laid down for the purpose.

The distribution of villages by different population size categories varies widely by states and Union Territories of the country (Table 1). In Himachal Pradesh, Uttarakhand, Arunachal Pradesh, Meghalaya and Andaman and Nicobar Islands, more than 40 per cent of the villages listed at the time of the 2011 population census were very small villages with a population less than or equal to 200 inhabitants at the time of the census. On the other hand, almost 93 per cent of the villages in Kerala had more than 5000 population at the 2011 population census. Kerala stands apart from rest of the states and Union Territories of the country in this regard as there is no other state in the country where the proportion of very large villages - villages with a population of more than 5000 - to the total number of villages listed at the 2011 population census exceeded 15 per cent with the only exception of Tripura (17 per cent). On the other hand, there is only one village out of 5258 villages in Arunachal Pradesh and 12 villages out of 17882 villages in Himachal Pradesh which had more than 5000 at the 2011 population census. In the Union Territory of Chandigarh, 3 out of 5 villages had more than 5000 population. Similarly, in the Union Territory of Puducherry, 29 out of 90 villages had more than 5000 population.

INDICATORS OF VILLAGE DEMOGRAPHIC SITUATION

Estimation of fertility, mortality and migration that are conventionally used to describe the demographic situation is not possible from PCA. It is however possible to estimate the following indicators on the basis of which, the village demographic situation can be characterised:

1. Ratio of the population aged 0-6 years to the population aged 7 years and above (ASI). This ratio gives an idea of the age structure of the village population. It gives an idea of the stage of demographic transition. It is well known that when this ratio is high, the population is at early stages of demographic transition. This ratio may be termed as the age structure index.

2. Ratio of the population aged 0-6 years to the female population aged 7 years and above (FTI). This ratio is very similar to the conventional child-woman ratio and gives an idea of the level of fertility. This ratio may be termed as the fertility index.

3. Ratio of the Scheduled Castes population to the total population (PSC). This ratio reflects the social class structure of the population. Scheduled Castes and Scheduled Tribes are official designations given to various groups of historically disadvantaged people.

4. Ratio of the Scheduled Tribes population to the total population (PST). This ratio, along with Scheduled Castes also reflects the social class structure of the population.

5. Ratio of the illiterate population to the total population aged 7 years and above (ILT). An illiterate in the Indian population census is a person at least 6 years of age who cannot read and write with understanding. This ratio reflects the level of literacy in the population. Higher is this index, the lower is the level of literacy in the population.

6. Ratio of workers to the total population aged 7 years and above (PWR). Work, in the Indian population census, is defined as participation of a person of any sex and age in any economically productive activity with or without compensation, wages or profit. Participation may be physical and/or mental. Work involves not only actual work but also supervision and direction. Part time help or unpaid work on farm, family enterprise or in any other economic activity is also classified as work. All persons

engaged in some type of work are classified as workers. This ratio is an indicator of the participation of the population in the village level social and economic production system. Higher is this ratio, the higher is the participation in the village social and economic production system. This ratio may be termed as the worker index.

7. Ratio of main workers to total workers (PMW). This ratio is an indicator of the opportunities available for participation in the village social and economic production system. All workers who have worked for at least six months during the year preceding the census are classified as main workers. Rest of the workers are classified as marginal workers.

The seven indicators described above may be termed as crude in their construct to describe village the demographic situation. The present analysis however shows that even these crude indicators are adequate to identify the distinguishing characteristics of demographic situation at the village level.

Estimation of the seven indicators described above for each of the 597619 inhabited villages of the country however revealed that, in many villages, estimates of these indicators appear implausible at the very first instance. One reason for this implausibility may be the composition of the population of the village concerned. For example, there were 18940 (2.95 per cent) villages where total number of households was less than 10. Similarly, there were 829 (0.13 per cent) villages where there was no female population and in 4441 (0.69 per cent) villages, there was no population aged 0-6 years. Moreover, there were 9062 (1.41 per cent) villages where the estimate of either ILT or PWR or PMW was found to be more than 1 which was not possible by definition. All these villages were, therefore, excluded from the analysis. At the same time, it was also observed that there were a number of villages where estimates of ASI and FTI were either extremely high or extremely low. As such, an exploratory data analysis exercise was carried out to identify outlier villages with respect to ASI and FTI. A village was classified as an outlier village if the estimate of either ASI or FTI for that village was found to be outside the range (Median $\pm$1.5*IQR). The exploratory data analysis revealed that there were 17594 (2.74 per cent) outlier villages in the country. These villages were also excluded from the analysis so that the present analysis is limited to 570959 (89.1 per cent) villages of the country.

Figure 1 presents kernel density plot (Parzen, 1962; Rosenblatt, 1956; Royal Society of Chemistry, 2006) of the distribution of the seven demographic

indicators across 570959 villages of the country estimated from the data available through PCA of 2011 population census. Summary measures of the inter-village distribution of the seven demographic indicators are given in table 2. The inter-village distribution of ASI, FTI, ILT and PWR is very close to the normal distribution which suggests that variation in ASI, FTI, ILT and PWR across the villages of the country may be termed as random. This randomness is also reflected from the very small difference between mean and median values of the index and the low coefficient of skewness. On the other hand, the inter-village distribution of PSC and PST is very highly positively skewed while that of PMW is very highly negatively skewed as is supported from the values of the coefficient of skewness and large difference between mean and median values of the index. A skewed inter-village distribution suggests that both Scheduled Castes and Scheduled Tribes population is concentrated in only a proportion of villages in the country whereas the proportion of main workers to total workers is generally low in majority of the villages but high to very high in some selected villages. This essentially means that the work force in the villages is largely constituted by marginal workers - workers who dis not worked for at least 180 days during the year preceding the population census. Lastly, the excess kurtosis is estimated to be negative in case of ASI, FTI, ILT, PWR and PMW but positive in case of PSC and PST which means that, except in case of the proportion of Scheduled Castes and Scheduled Tribes population, the inter-village distribution of the remaining five indicators is essentially platykurtic in shape characterised by lower, flat peak and corresponding thin tails so that villages are less clustered around the central tendency. In case of the proportion of Scheduled Castes and Scheduled Tribes, the inter-village distribution is leptokurtic in shape characterised by long, high peak and corresponding flat tails so that villages are heavily clustered around the central tendency.

DOMAINS OF DEMOGRAPHIC SITUATION

The first step in patterning demographic regime was devoted to identifying domains of village demographic situation on the basis of the seven indicators. Factor analysis technique (Sharma, 1996) was applied for the purpose which combined seven indicators into domains of demographic situation in such a way that indicators within a domain were highly correlated whereas indicators in different domains were not correlated. Factor analysis also permitted constructing composite index for each domain.

Application of the factor analysis technique revealed that the seven demographic indicators can be grouped into three factors or domains of the demographic situation and the three domains accounted for almost 75 per cent of the total variance in the original data set. The KMO measure of sampling adequacy was found to be 0.661 reflecting the homogeneity of the indicators. This means that partial correlations among the indicators are small and patterns of correlations are relatively compact. On the other hand, Bartlett's test for sphericity was found to be statistically significant suggesting that the correlation matrix was not the identity matrix and that the correlations in the original data set were appropriate for factor analysis. Both these tests confirmed that the suitability and appropriateness of factor analysis procedure to combine the seven demographic indicators.

The demographic domains revealed through the application of factor analysis is presented in table 3 along with the proportion of the variance explained by each domain. The table shows only those indicators which have a factor loading of at least 0.60. The first demographic domain has high loadings in three indicators - age index (ASI), fertility index (FTI) and literacy index (ILT). This domain accounts for more than 37 per cent of the total variance. Indexes ASI and FTI reflect fertility and mortality levels in the population. The structure of this domain confirms that fertility and mortality level at the village level is strongly associated with the level of literacy - the higher is the index ILT, the higher are the indexes ASI and FTI.

The second demographic domain has high loadings in two indicators - proportion of Scheduled Castes population (PSC) and proportion of Scheduled Tribes population (PST). This demographic domain accounts for more than 20 per cent of the total variance. This domain reflects the social class composition of the village population and, therefore reflects the social and cultural dimension of the village demographic situation.

Finally, the third demographic domain has high loadings in the ratio of total workers to the population aged 7 years and above (PWR) and the ratio of main workers to total workers (PMW). This domain accounts for more than 17 per cent of the total variance. This domain reflects the extent and the nature of the participation of the village people in the village level social and economic production system.

Application of the factor analysis procedure thus suggests that village demographic situation can be described in terms of three demographic domains

which are mutually independent. On the basis of the result of the factor analysis, domain index have been calculated for the three domains for each village following the method proposed by Nicolette and others (2000). These indexes can be used to describe the demographic situation of any village on a three dimensional demographic space. This demographic space comprises of the domain characterised by levels of fertility, mortality and literacy; domain characterised by social class composition of the population; and domain characterised by participation in the village social and economic production system.

CLUSTERING OF VILLAGES

Factor analysis also provided domain score for each village in each of the three domains of the demographic space. These scores were used to cluster villages in terms of the three demographic domains. The TwoStep cluster approach (SPSS, 2001) was applied for the purpose. In the first step of this approach sub-clusters are constructed while are grouped into clusters in the second step by integrating hierarchical and partitioning clustering. This approach automatically determines the optimal number of clusters and permits identification of villages which do not fit in any of the identified clusters.

Cluster analysis revealed that 570959 villages of the country may be grouped into six clusters on the basis of the seven indicators. The first cluster comprises of 115369 (20.2 per cent) villages; the second cluster 107303 (18.8 per cent) villages; the third cluster 62117 (10.9 per cent) villages; the fourth cluster 137838 (24.1 per cent) villages; the fifth cluster 64836 (11.4 per cent); and the sixth cluster 83332 (14.6 per cent) villages of the country. In addition, there are 164 outlier villages in the sense that they did not fit into any of the six clusters. The average Silhoutte measure of cohesion and separation was found to be 0.3 which means that cohesion and separation between the six clusters are fair and therefore acceptable. Discriminant analysis revealed that the clustering exercise correctly classified 86.4 per cent of the villages of the country included in the analysis. The correct classification rate was 95.4 per cent for villages of the first cluster; 93.1 per cent for villages of the second cluster; 94.5 per cent for villages of the third cluster; 79.9 per cent for villages of the fourth cluster; 91.3 percent for villages of the fourth cluster; 91.3 per cent for villages of the fifth cluster; and 66.3 per cent for villages of the sixth cluster. Clustering exercise thus revealed six distinct demographic regimes in the country.

Table 4 summarises the demographic scenario in six cluster of villages in terms of a number of demographic indicators including the seven indicators used for clustering purpose. Villages in the first cluster account for 23.5 per cent of the rural population of the country whereas villages in the second cluster account for 22.8 per cent of the rural population. Similarly, villages in the third cluster account for 15 per cent of the rural population; villages in the fourth cluster account for 23.8 per cent of the rural population; villages in the fifth cluster account for 5.7 per cent of the rural population; and villages in the sixth cluster account for 9.1 per cent of the rural population of the country. The proportionate distribution of households in each cluster is different from the proportionate distribution of population so that the average household size varies by cluster. The average household size is the largest in cluster 3 but the smallest in cluster 1. In three clusters, the average household size is more than 5 persons per household which is well above the national average.

Table 4 shows that the demographic situation in the six clusters is essentially different. Indexes ASI and FTI are the lowest in cluster 1 but the highest in cluster 3. This implies that demographic transition in villages of cluster 1 is relatively at an advanced stage whereas demographic transition in cluster 3 is relatively at an early stage. In cluster 5 also, the demographic transition is at a relatively early stage. On the other hand, the level of literacy is relatively the highest in cluster 1 but relatively the lowest in cluster 5. Similarly, the proportion of Scheduled Castes population is very high in cluster six whereas the proportion of scheduled tribes population is very high in cluster 5 and in cluster 6. Finally, PWR is found to be relatively high in clusters 5 and 6 but relatively low in clusters 2 and 3. PMW, on the other hand, is found to be relatively high in clusters 1 and 2 but relatively low in cluster 4.

From the perspective of population and development, the situation in villages of cluster 1 appears to be relatively the best whereas the situation in villages of cluster 3 appears to be relatively the poorest. Villages of cluster 1 appears to be relatively an advanced stage of demographic transition. The demographic situation in these villages is associated with highest level of literacy and a very high proportion of main workers to total workers which reflect ample opportunities of participation in the village level social and economic production system.

On the other hand, villages of cluster 3 appear to be relatively the lowest stage of demographic transition and the demographic situation in villages of this cluster is associated with very low level of literacy and lowest level of the

participation in the village level social and economic production system. The proportion of main workers to total workers, however, is somewhat better in villages of this cluster. A similar population and development situation appears to prevail in villages of cluster 5 also where both the level of literacy and the proportion of main workers to the total workers is relatively the lowest in the country. Finally, the demographic and development scenario in villages of cluster 2 is relatively better than that in villages of cluster 4.

Table 5 shows the distribution of villages of different clusters across state/Union Territory of the country. Although, the distribution of villages by clusters is different in different states and Union Territories, yet regional pattern in the distribution is very much apparent. For example, throughout the southern part of the country, the dominance of villages belonging to cluster 1 is very clear with the only exception of the Union Territory of Lakshadweep where all villages belong to cluster 6. Villages belonging to this cluster are also dominant in the northern states of the country, particularly in Himachal Pradesh, Uttarakhand and Punjab. On the other hand, villages belonging to cluster 3 and cluster 4 are dominant in Bihar while villages belonging to cluster 5 and cluster 6 are dominant in most of the north-eastern states. Another interesting observation of table 5 is that throughout central India and in north-eastern states, less than 10 per cent of the villages belong to cluster 1. The cluster pattern is however not so contrasting in other states of the country such as Jammu and Kashmir, Uttar Pradesh, Madhya Pradesh and Gujarat. In these states, villages are distributed across different clusters which suggest that different demographic regimes prevail in different parts of these states.

The analysis suggests that there are at least six demographic regimes that prevail. The distinguishing feature of the first demographic regime is very low fertility and mortality and very high level of literacy. Another distinguishing features of this regime is very high proportion of main workers in the work force. This demographic regime is prevalent mostly in the southern and northern regions of the country.

The second demographic regime, on the other hand, is characterised by moderately low levels of fertility and mortality and average level of literacy. The distinguishing features of this demographic regime include a large share of marginal workers in the work force and a large proportion of Scheduled Tribes in the population. This regime is primarily prevalent in the north-eastern part of the country.

The third demographic regime is characterised by average levels of fertility and mortality and literacy and low level of participation in the social and economic production system, although very high proportion of workers in this regime are main workers. This demographic regime is particularly prevalence in Delhi, Haryana, Punjab and in Union Territories of Chandigarh and Daman and Diu. Presence of this regime may also be seen in Gujarat, and Uttar Pradesh.

In the fourth demographic regime, level of fertility, mortality and literacy is very close to national average. The distinguishing feature of this regime is very low proportion of main workers to total workers and very high proportion of Scheduled Castes. This regime is largely prevalent in northern and central India.

The fifth demographic regime is characterised by very high levels of fertility and mortality coupled with very low level of literacy. The distinguishing feature of this regime is that very close of 80 per cent of the population under this regime is Scheduled Tribes. This regime is prevalent in the north-eastern and central region of the country.

The sixth and the last demographic regime is characterised by the highest levels of fertility and mortality along with very low level of literacy and lowest level of participation in the social and economic production system. The proportion of main workers to total workers is marginally above the national average in this regime. A distinguish feature of this demographic regime is very low proportion of Scheduled Tribes population. The dominance of this demographic regime is confined to only three states - Bihar, Jammu and Kashmir and Uttar Pradesh.

Conclusions

This paper has attempted to identify different demographic regimes that prevail in villages of India. The analysis reveals that the villages of India can be grouped into six demographic regimes with distinguished demographic characteristics. The analysis also reveals that different demographic regimes have strong regional pattern. This is so when the social, cultural, economic and environmental diversity of India is well known. The present analysis shows that despite this diversity, the village demographic situation can be described in terms of a small number of demographic regimes. These similarities in village demography have important implications to population and development planning and programming at the grass roots level.

References

Davis K (1945) The world demographic transition. *Annals of the American Academy of Political and Social Science* 273(1):1-11. doi:10.1177/000271624523700102.

Demeny P (1972). Early fertility decline in Austria-Hungary: A lesson in demographic transition. In: DV Glass and R. Revelle (eds) *Population and Social Change*. London, Edward Arnold.

Government of India (2011) *Census of India 2011. Provisional Population Totals. Rural-Urban Distribution. India, Series 1*. New Delhi, Registrar General and Census Commissioner of India.

Landry A (1934) *La Revolution Demographique*. Paris, Librarie due Recusil Sirey.

Nicolette G, Scarpetta D, Boylaud O (2000) Summary indicators of product market regulation with an extension to employment protection legislation. OECD, OECD Economics Department Working Paper 236.

Notestein FW (1945) Population: The long view. In TW Schulz (ed.). *Food for the World*. Chicago: Chicago University Press.

Parzen E (1962) On estimation of a probability density function and mode. *Annals of Mathematical Statistics* 33(3):1065-76.

Pressat R (1985) *The Dictionary of Demography*. Oxford, Basil Blackwell.

Rosenblatt S (1956) Remarks on some non-parametric estimates of a density function. *Annals of Mathematical Statistics* 27(3):832.

Royal Society of Chemistry (2006) Representing data distributions with kernel density estimates. *AMC Technical Brief No. 4*. London, Statistical Sub-committee.

Sharma S (1996) *Applied Multivariate Techniques*. New York, John Wiley and Sons.

SPSS (2001) *The SPSS TwoStep Cluster Component.* SPSS Inc.

Table 1
Distribution of villages by population size in states and Union Territories, 2011

India/State/Union Territory	Population of the village								
	≤200	201-500	501-1000	1001-2000	2001-3000	3001-4000	4001-5000	>5000	Total
Jammu and Kashmir	521	1273	1520	1637	678	320	153	235	6337
Himachal Pradesh	8535	5847	2439	829	171	36	13	12	17882
Punjab	971	1902	3235	3469	1323	594	303	371	12168
Chandigarh	0	0	0	0	0	1	1	3	5
Uttarakhand	7873	4649	1813	823	271	128	71	117	15745
Haryana	233	440	1034	1967	1142	702	409	715	6642
Delhi	4	2	6	21	23	22	10	15	103
Rajasthan	4733	8897	12393	10498	3582	1481	730	950	43264
Uttar Pradesh	7917	13609	23361	27977	12458	5648	2871	3973	97814
Bihar	2437	4589	7523	10068	5349	3034	1730	4343	39073
Sikkim	27	78	147	127	32	9	1	4	425
Arunachal Pradesh	3791	1014	320	111	18	1	2	1	5258
Nagaland	215	421	336	239	99	40	19	31	1400
Manipur	807	774	412	259	108	64	45	46	2515
Mizoram	87	236	231	105	32	9	2	2	704
Tripura	11	52	80	197	170	118	87	148	863
Meghalaya	2671	2503	943	253	56	21	5	7	6459
Assam	4343	5639	6062	5702	2029	893	376	328	25372
West Bengal	3684	6847	8554	8725	4101	2198	1240	2129	37478

India/State/Union Territory	Population of the village								
	≤200	201-500	501-1000	1001-2000	2001-3000	3001-4000	4001-5000	>5000	Total
Jharkhand	5511	8255	7907	5202	1509	596	253	259	29492
Orissa	11370	13257	11925	7872	2090	700	259	204	47677
Chhattisgarh	1798	4264	6135	5374	1310	437	159	90	19567
Madhya Pradesh	5490	11969	16304	12478	3414	1205	456	613	51929
Gujarat	583	1908	3894	5557	2755	1326	693	1127	17843
Daman and Diu	0	2	3	5	1	1	2	5	19
Dadra and Nagar Haveli	1	5	6	18	7	15	6	7	65
Maharashtra	2619	6427	11143	12134	4442	1870	839	1485	40959
Andhra Pradesh	2859	3244	4243	6388	3754	2118	1276	2404	26286
Karnataka	3166	5300	7026	6479	2584	1242	604	996	27397
Goa	30	57	56	75	44	33	10	15	320
Lakshadweep	2	1	0	0	1	1	0	1	6
Kerala	1	1	3	11	17	23	18	943	1017
Tamil Nadu	479	1171	2380	4222	2574	1567	958	1698	15049
Puducherry	0	0	2	13	21	12	13	29	90
Andaman & Nicobar Islands	178	78	64	50	12	9	3	2	396
India	82947	114711	141500	138885	56177	26474	13617	23308	597619

Source: Census 2011.

Table 2
Summary measures of inter-village distribution of demographic indicators

Measure	ASI	FTI	ILT	PSC	PST	PWR	PMW
Minimum	0.027	0.037	0.028	0.000	0.000	0.002	0.000
Q1	0.114	0.262	0.372	0.007	0.000	0.389	0.489
Median	0.141	0.335	0.480	0.119	0.001	0.522	0.738
Q3	0.170	0.423	0.604	0.270	0.196	0.634	0.924
Maximum	0.262	0.687	1.000	1.000	1.000	1.000	1.000
Mean	0.144	0.348	0.495	0.179	0.184	0.518	0.678
SD	0.039	0.114	0.171	0.205	0.323	0.153	0.275
Skewness	0.211	0.430	0.409	1.605	1.640	0.128	-0.698
Kurtosis	-0.316	-0.228	-0.148	2.743	1.130	-0.509	-0.519
N	570959	570959	570959	570959	570959	570959	570959

Source: Author's calculations.

Table 3
Factor structure

Factor I	Factor II	Factor III
1. Age index (ASI) 2. Fertility index (FTI) 3. Literacy index (ILT)	1. Scheduled Castes index (PSC) 2. Scheduled Tribes index (PST)	1. Worker index (WPR) 2. Main worker index (MPW)
Sums of squared loadings		
2.607	1.419	1.197
Proportion of total variance explained		
37.25	20.26	17.01

Remark: Only those indicators are shown in each factor which have a factor loading of more than 0.60.

Table 4

Demographic situation different clusters

Indicator	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Total
Distribution of villages, households and population							
Villages	20.21	18.79	10.88	24.14	11.36	14.62	100.00
Households	27.25	21.62	13.27	22.32	5.67	9.87	100.00
Population	23.52	22.84	15.02	23.76	5.73	9.11	100.00
Population (0-6)	16.60	22.45	20.36	25.37	7.26	7.96	100.00
Scheduled Castes	22.25	18.48	10.70	43.71	0.93	3.91	100.00
Scheduled Tribes	9.44	7.73	3.49	6.06	41.59	31.69	100.00
Literates	28.18	24.04	11.57	22.68	4.29	9.24	100.00
Workers	25.74	19.68	11.94	23.55	7.04	12.05	100.00
Main workers	29.95	23.58	12.25	17.84	5.83	10.57	100.00
Marginal workers	15.61	10.31	11.20	37.30	9.95	15.61	100.00
Estimates of demographic indicators							
ASI	0.1134	0.1652	0.2431	0.1820	0.2232	0.1440	0.1686
FTI	0.2287	0.3413	0.5087	0.3753	0.4483	0.2905	0.3453
ILT	0.2223	0.2850	0.4416	0.3423	0.4664	0.3239	0.3187
PSC	0.1859	0.1505	0.1326	0.3421	0.0300	0.0798	0.1860
PST	0.0435	0.0367	0.0252	0.0276	0.7862	0.3770	0.1084
PWR	0.5093	0.4197	0.4133	0.4898	0.6276	0.6324	0.4886
PMW	0.8219	0.8461	0.7244	0.5348	0.5846	0.6195	0.7063
N	115369	107303	62117	137838	64836	83332	570795

Source: Author's calculations

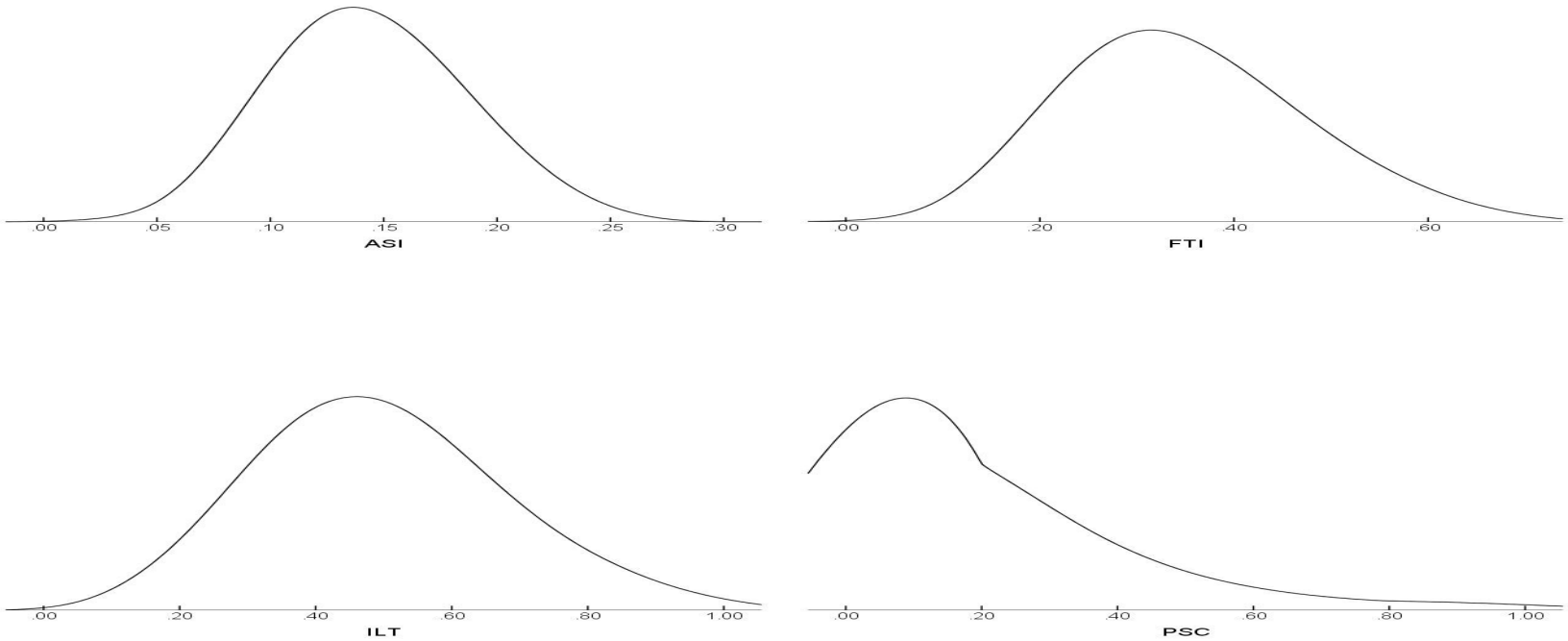
Table 5
Distribution of villages by clusters across states/Union Territories

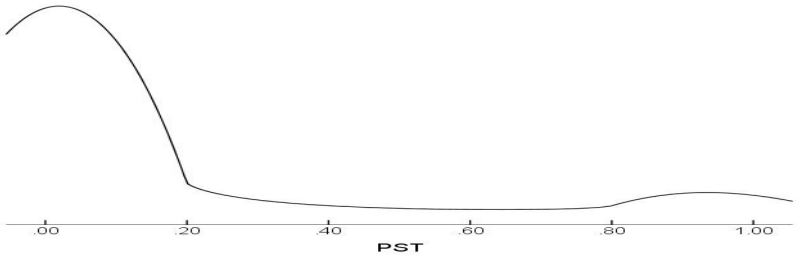
State/UT	Cluster					
	1	2	3	4	5	6
North India						
Jammu and Kashmir	7.4	25.9	23.5	24.9	9.3	9.1
Himachal Pradesh	35.7	7.2	0.8	25.7	1.6	29.1
Punjab	37.4	33.8	0.9	27.1	0.0	0.8
Chandigarh	0.0	100.0	0.0	0.0	0.0	0.0
Uttarakhand	30.7	20.9	4.5	27.3	1.1	15.5
Haryana	19.0	47.0	6.9	24.9	0.0	2.2
Delhi	13.7	80.4	2.0	2.9	0.0	1.0
Central India						
Rajasthan	6.2	16.9	14.0	33.0	16.2	13.7
Uttar Pradesh	4.5	30.5	20.9	41.9	0.5	1.8
Bihar	1.3	15.2	44.7	35.7	1.9	1.2
Jharkhand	2.2	8.9	12.1	27.7	34.6	14.6
Chhattisgarh	7.7	11.3	2.6	15.0	28.4	35.1
Madhya Pradesh	7.4	19.5	8.6	24.6	22.3	17.7
East India						
Sikkim	29.1	4.9	0.0	0.2	3.7	62.1
Arunachal Pradesh	3.9	5.6	4.5	0.3	53.3	32.5
Nagaland	1.1	1.4	0.9	0.0	52.3	44.3
Manipur	8.4	8.0	1.8	2.2	30.9	48.7

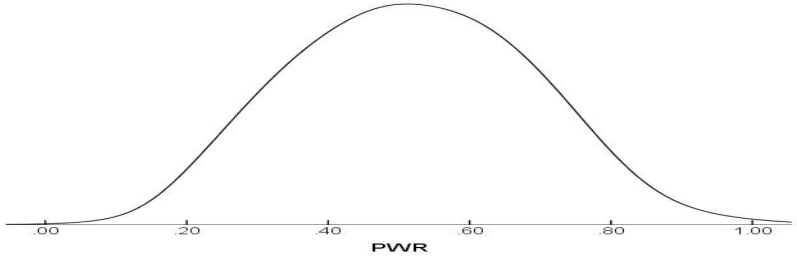
State/UT	Cluster					
	1	2	3	4	5	6
Mizoram	0.2	1.1	2.1	0.0	65.2	31.5
Tripura	24.4	14.0	0.1	7.1	24.7	29.6
Meghalaya	1.1	2.8	4.9	0.4	77.1	13.8
Assam	19.4	20.8	18.4	8.1	13.1	20.3
West Bengal	29.9	17.5	3.6	32.3	3.3	13.4
Orissa	25.9	11.7	1.4	16.6	20.8	23.7
West						
Gujarat	24.8	28.0	6.8	4.5	13.8	22.1
Daman and Diu	47.4	36.8	0.0	0.0	0.0	15.8
Dadra and Nagar Haveli	1.5	1.5	0.0	0.0	73.9	23.1
South						
Maharashtra	45.0	21.1	1.1	4.0	7.9	20.9
Andhra Pradesh	42.6	10.9	0.6	11.5	10.0	24.4
Karnataka	50.3	14.8	2.1	15.6	2.1	15.1
Goa	82.8	4.7	0.0	0.0	0.9	11.6
Lakshadweep	0.0	0.0	0.0	0.0	0.0	100.0
Kerala	85.3	14.0	0.0	0.0	0.0	0.8
Tamil Nadu	64.3	6.2	0.1	21.0	1.0	7.4
Puducherry	66.7	17.8	0.0	15.6	0.0	0.0
Andaman & Nicobar Islands	53.2	21.3	1.5	1.5	10.0	12.5

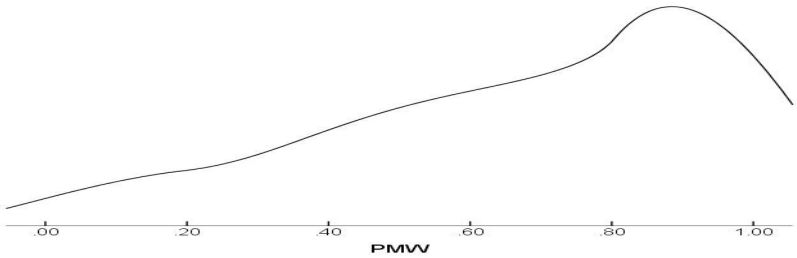
Source: Author's calculations

Figure 1
Kernel density plots of demographic variables



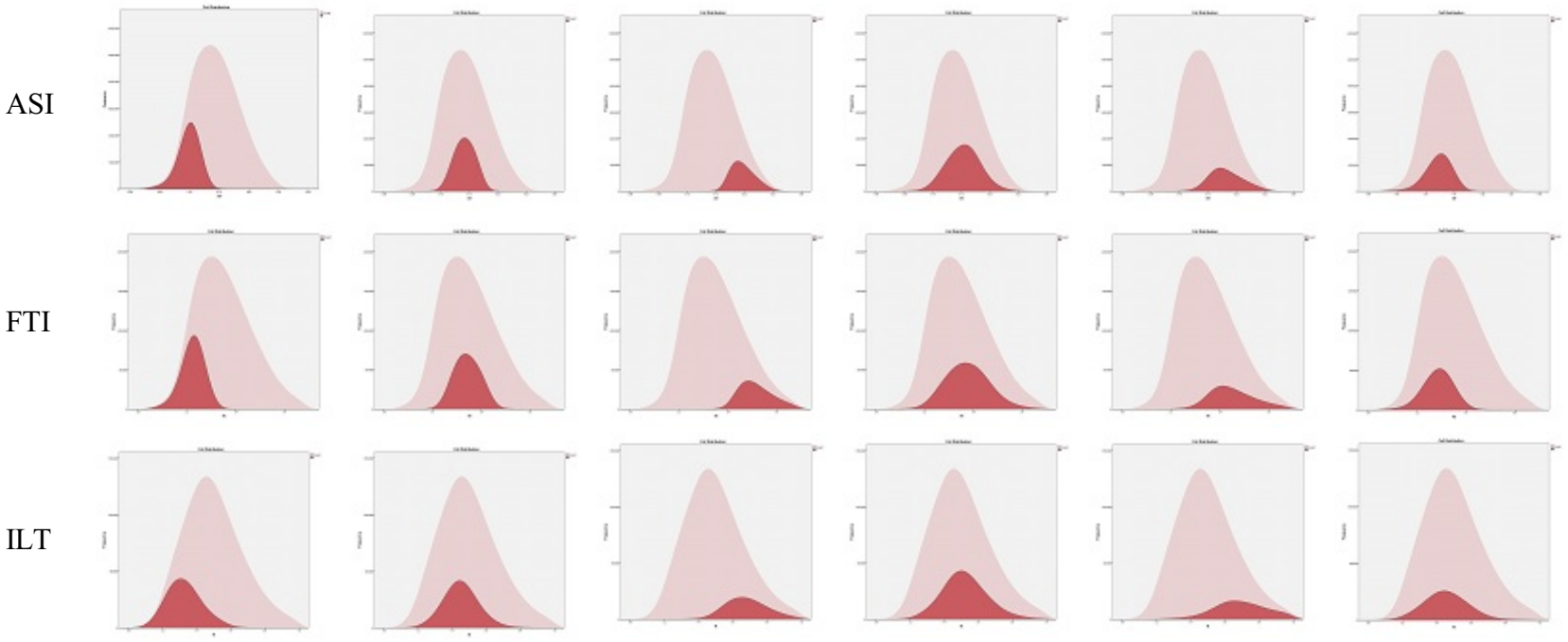


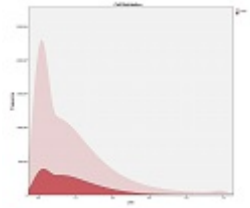
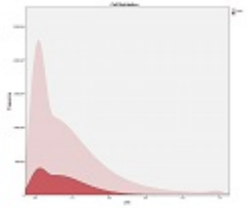
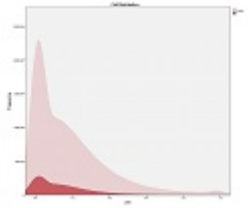
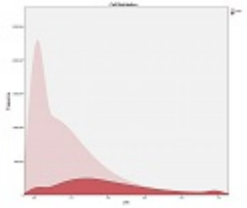
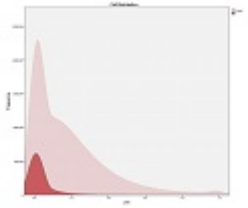
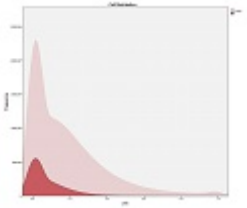
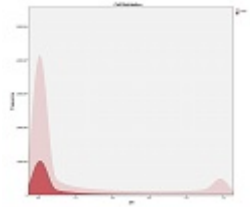
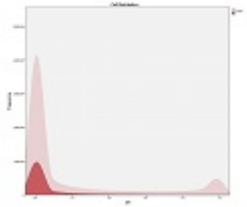
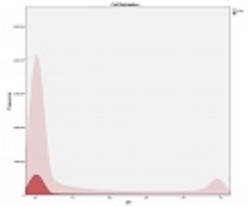
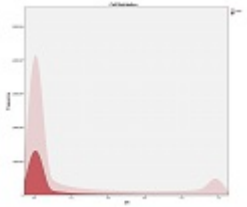
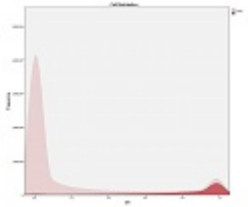
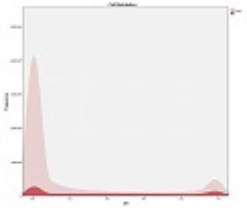
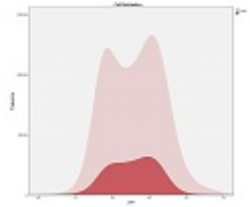
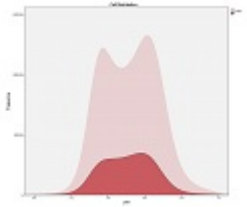
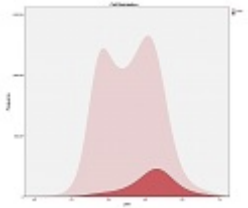
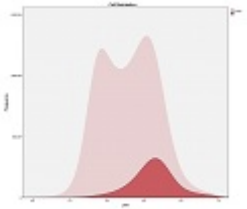
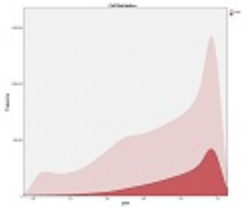
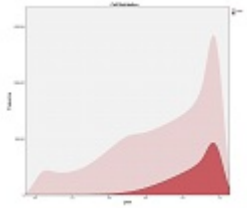
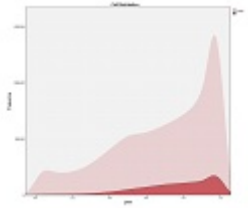
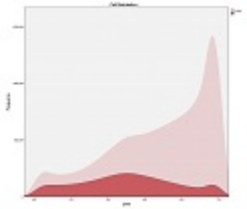
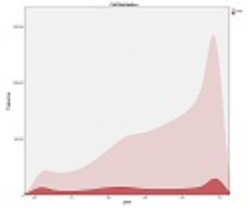
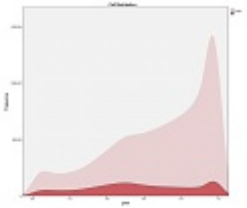




ASI	Age structure index
FTI	Fertility index
ILT	Literacy index
PSC	Proportion of Scheduled Castes
PST	Proportion of Scheduled Tribes
PWR	Worker index
PMW	Main worker index

Figure 2
Distribution of villages in different clusters



	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
PSC						
PST						
PWR						
MPW						

Source: Author's calculations