Demography India (2017) ISSN: 0970-454X
Vol. 46, Issue: 2, pp: 14-30

Research Article

Population and Development Morphology of Villages in India

Aalok Ranjan Chaurasia[1,*]

Abstract

This paper uses data from India's 2011 population census to analyse population and development morphology of villages in India. The village population and development morphology is captured through a composite population and development index that has been developed for the purpose. The analysis suggests that the villages in India can be grouped into eight clusters having distinct population and development morphology. The analysis reveals that the population and development scenario in the village is strongly influenced by the village population structure characterised in terms of gender balance and social class composition. Moreover, while the demographic transition in the village appears to be linked with the level of literacy in the village, it has no link with the level of participation in the productive activities. The paper emphasises the need of a spatial approach to population and development planning and programming with the village as the basic planning and implementation unit.

Introduction

The aim of this paper is to explore the population and development morphology of villages in India. Population and development morphology may be conceptualised on the lines of social morphology which has been described as the study of forms and structures of the society (Durkheim, 1982). According to Davis (1955), social morphology can be developed in terms of 'structure', 'process' and 'stage' of any social phenomenon. Accordingly, the population and development morphology may be developed in terms of the 'structure' of the population, the 'state' of development and the 'stage' of population transition. The 'structure' of the population may be perceived as a sociological perspective; the 'state' of development as a human welfare perspective; and the 'stage' of population transition as a demographic perspective. The population and development morphology is then concerned with the characterising population and development landscape in terms of differences in the structure of this landscape. Following the Durkheim terminology, the structure of population and development landscape comprises of a base and a super-structure. The base of the population and development landscape is the population characteristics whereas the super-structure is the development activities which essentially originate from the base. Finally, the differences in the population and development landscape may be analysed across spatial units such as village which may be perceived as the visible manifestation of the population and development landscape. In this conceptualisation, the population and development morphology also depicts the relationship between the population dimension and the development dimension of the population and development landscape and how this relationship varies across spatial units.

This paper has three objectives. The first objective of the paper is to construct a composite population and development index to reflect the population and development landscape. The second objective is to estimate the composite population and development index to characterise the population and development landscape in more than 517 thousand villages of the country and analysing the variation in the population and development landscape across villages. The third and the last objective of the paper is to apply data mining tools to explore how the population and development landscape of a village is conditioned by the defining characteristics of the village. The

[1]Professor, MLC Foundation and 'Shyam' Institute, Bhopal, Madhya Pradesh
Email: aranjan@shyaminstitute.in

study of the population and development morphology of villages matters for India as almost 70 per cent of India's population lives in rural areas distributed across more than 640 thousand villages of varying population size (Government of India, 2011). Villages in India have always been an integral part of the agricultural, industrial and commercial landscape of the country. People living in villages are conditioned by a very diverse and heterogeneous set of social, religious, cultural, natural and economic environment that shapes their capabilities and decides their participation in the social and economic production system. It is well-known that the population and development landscape of the country India is defined, to a significant extent, by the population and development landscape of Indian villages.

Another reason for studying the population and development morphology of Indian villages is the recent shift in the official approach to well-being and welfare of the rural people from rural development to village development. A number of schemes have recently been launched in the country and in its constituent states to set the trend in this direction. Rural development in India has traditionally followed the sectoral approach in which government development policies and programmes are conceived and organised along different development sectors with little or limited integration among them. The village development approach, on the other hand, is essentially a spatial approach that addresses village specific population and development issues and concerns in an integrated manner. The shift to village development approach, therefore, requires characterisation of the population and development landscape and analysing how this landscape varies across villages. The present paper is an attempt in this direction.

Very little is currently known about the distinguishing features of population and development landscape in Indian villages. Villages, in India, have been studied extensively through sociological and anthropological perspectives since the colonial period. These studies have focussed primarily on the structure of the village, especially, in the context of social class (Beteille, 1969). The underlying argument of these studies is that social class or caste made the village a social reality (Srinivas, 1952; Bailey, 1957, 1963; Beteille, 1965; Kessinger, 1971; Shah, 1973; Chakravarty, 1975). Village, in these studies, has also been a suitable locale for the study of peasant society and culture (Redfield, 1955). However, village, as a population and development entity, has rarely been paid attention in either the development discourse or in the demographic research in India. This is so when issues of national development have been closely identified with the upliftment of villages even before the independence (Gandhi, 1944; Nehru, 1961) and projection of the village as a template for nation building after independence (Thakur, 2014). There are some studies that have attempted to analyse the relationship of village characteristics with the use of maternal and child health services (Ghosh and Singh, 2004; McNay, 2002; Stephenson and Tsui, 2002). In a recent study, it has been observed that the size of the village population matters as large villages are found to be relatively better developed than small or medium size villages (Singh, Chakraborty and Roy, 2008). The study, however, considers only the size of the village population in exploring the village population and development morphology.

The study of village population and development morphology also contributes to the long and enduring debate on population and development interrelationship. The complexity of this relationship is well reflected in frequent oscillations in the scientific wisdom about macroeconomic consequences of population growth. Three alternative positions define these oscillations: population growth either restricts or promotes or is independent of economic growth (Bloom, Canning and Sevilla, 2001). There are studies that have attempted to explore population and development relationship at micro - household - level (Sinding, 2009) but there is no study, especially in India, that has attempted to study the population and development morphology at meso - village – level which is the lowest spatial unit at which the population dimension interacts with the development dimension of the population and development landscape.

The paper is organised as follows. The next section of the paper discusses the concept of the village as adopted in the population census in India. The village, in the Indian population census, is essentially an administrative unit with clearly marked geographical boundaries rather than a human settlement. A village in India may comprise of more than one human settlements of varying population size. Section three describes the data source pertaining to village level population and development landscape in India. Although, limited in scope, the only source of village level population and development related data in India is the decennial population census. Section four of the paper outlines the construction of the population and development index that serves the basis for the study of village population and development morphology. The fifth section presents an analytical perspective of the population and development morphology across villages in India whereas the sixth section interrelates the village population and development landscape with selected defining characteristics of the village. The seventh and the last section of the paper summarises the findings and discusses their policy implications.

Village in the Indian Population Census

The concept of the village adopted in the decennial population census in India is different from the conventional concept of village as a human settlement. A village is conventionally defined as a human settlement which is larger than a hamlet and smaller than a town; a hamlet has a tiny population less than 100 (Doxiadis, 1968). In the population census in India, however, a village is defined as an administrative unit with non-overlapping geographical boundaries. The entire geographical area of the country is first divided into urban and rural areas. Urban areas are defined according to a clearly articulated definition. The population living in the urban areas is classified as the urban population while the population not residing in the urban areas is classified as the rural population and is organised into administrative areas following the administrative boundaries of revenue villages (Government of India, 2011a). These administrative areas are termed as villages in the Indian population census. A village, defined in this manner, is not a human settlement in the true sense but the lowest level administrative unit with well-defined administrative or geographical boundaries. There may be more than one human settlements or there may be no human settlement within the administrative boundary of a village. If there is no human settlement within the administrative boundaries of the village, then the population of the village is zero and the village is christened as an uninhabited village. On the other hand, if there are more than one human settlements within the administrative boundaries of the village, then the population of all human settlements within the village is added to obtain the village population, although, the population characteristics of different human settlements within the same village may be different. Moreover, no attention is paid to the permanent or the temporary nature of human settlements within the administrative boundaries of the village. There may be only permanent or only temporary human settlements or both within a village. This conceptualisation of defining the village for the purpose of population census, has implications for studying the population and development morphology. There may be a possibility that the population and development landscape of different human settlements within the same village may be substantially different and the population and development landscape of the village may not reflect the population and development landscape of any of the human settlements that are encompassed in the village. The Census Commissioner of India, however, prepares the list of all human settlements – permanent or temporary – within the administrative boundaries of a village before every decennial population census. An examination of this list suggests that the total number of human settlements in the country is significantly higher than the total number of villages as identified for the purpose of population census. On the other hand, although, population of all human settlements is enumerated at the time of the decennial population census, yet, human settlement-specific data are not released by the Census Commissioner to allow for characterising population and development landscape of each human settlement. The data collected at the population census are available only for villages irrespective of the number of human settlements within the administrative boundaries of the village. As such, the village is the lowest administrative unit for the study of population and development morphology in India.

Data

The present analysis is based on the data available through the primary census abstract of India's 2011 population census (PCA 2011). The primary census abstract is the only source of population and development related data at the village level in India. The PCA 2011 provides data about total number of households; total population; population below 7 years of age along with the social class composition of the population for every village of the country as identified at the time of the 2011 population census. PCA 2011 also provides data about the educational and work status of the population. Workers are classified into main and marginal workers. Main workers are those who have worked at least six months during the year prior to the census. All other workers are classified as marginal workers. A worker, either main or marginal, is further classified into one of the four occupational categories - cultivator; agricultural labourer; household industry worker and other worker (Government of India, 2011).

Table 1: Distribution of villages in India by population size, 2011

Population	Number of villages according to 2011 population census		Number of villages included in the present	
	Number	Percent	Number	Percent
Uninhabited	43330	6.8	0	0.0
< 100	40958	6.4	0	0.0
100 - 500	156700	24.4	131481	25.4
500-1000	141500	22.1	133637	25.8
1000-3000	195062	30.4	189915	36.7
3000-5000	40091	6.3	39367	7.6
≥ 5000	23308	3.6	22905	4.4
Total	640949	100.0	517305	100.0
Median	751		970	
IQR	1312		1367	

Source: Author’s calculations

There were 640949 villages of varying population size in India listed at the 2011 population census, out of which 43330 (6.8 per cent) were uninhabited (Table 1). The population of the remaining 597619 (93.2 per cent) villages ranged from 1 to 66062. There were 40958 (6.4 per cent) villages with a population less than 100. These villages were essentially hamlets according to the settlement hierarchy proposed by Doxiadis (1968). Therefore, 84288 villages having either no population or very small population at the 2011 population census were excluded from the analysis. An exploratory analysis of data pertaining to remaining 556661 villages, however revealed that indicators depicting village population and development landscape appeared inconsistent in 39356 villages so that so that the present analysis is restricted to 517305 (80.7 per cent) villages of the country. On the other hand, there were 22905 (3.6 per cent) villages having a population of at least 5000 at the 2011 population census but they could not be classified as an urban area because they did not conform fully to the definition of the urban area adopted at the 2011 population census.

Population and Development Index

Studying population and development morphology requires that the three components of the population and development landscape – ‘structure’, ‘state’ and ‘stage’ - are connected by assembling them under one rubric which may serve as an integrated representation of population and development landscape. This composite rubric also serves as the basis for describing how the population and development landscape varies across space. The first task in the study of village population and development morphology, therefore, is to quantify this composite rubric in the form of a composite population and development index that encompasses village population ‘structure’, ‘state’ of development in the village and the ‘stage’ of transition of the village population. This means that village level indicators depicting the population ‘structure’, the ‘state’ of development and the ‘stages’ of population transition are needed to construct the population and development index. The indicators may be selected on the basis of some theoretical construct or through a policy perspective or on the basis of data availability (Hanafin and Brooks, 2005; Chaurasia, 2016). Ideally, all three approaches should be taken into consideration while selecting the indicators (Bauer et al, 2003) but availability of the necessary data is generally the prime consideration for selecting indicators.

An examination of PCA 2011 data suggests that the following five indicators can be calculated for each of the 517305 villages of the country that have been included in the present analysis to reflect the population and development landscape of the village:

1. The proportion of the population of the village aged less than 7 years (*ASI*). This proportion reflects the age 'structure' of the village population, the higher the proportion the younger the village population.
2. The ratio of the population aged less than 7 years to the females age 7 years and above in the village (*FTI*). This ratio, essentially, reflects the level of fertility in the village and thus reflects the 'stage' of transition in the village population, the higher the ratio the earlier the stage of transition of the village population.
3. The proportion of the population aged 7 years and above who cannot read and write with understanding (*ILT*). This proportion reflects the 'state' of development in the village, the higher this proportion, the poorer the 'state' of development.
4. The proportion of non-workers in the village population (*NWR*). This proportion also reflects the 'state' of development in the village, the higher this proportion the poorer the 'state' of development in the village.
5. The proportion of marginal workers to the total workers in the village (*MAR*). This proportion also reflects the 'state' of development in the village, the higher this proportion the poorer the 'state' of development.

The rationale behind selecting these indicators can be justified conceptually. Although, the interaction between productive participation, education, disposable income, welfare expenditure and income inequality is quite complex, yet, it is well known that non-participation in productive activities reduces disposable income and increases income inequality. Non-participation in productive activities has also been associated with high crime rate, violence and social unrest. On the other hand, the age composition of the population reflects the stage of population transition as linkages between population transition and transition in the population age composition is well-known. At the early stages of transition, population remains young with the age pyramid typically triangular in shape with broad base and thin top. As population transition progresses, the population gets older and the shape of the population pyramid changes from a triangular one to a rectangular one. At the end of the population transition, there is a heavy concentration of the population in older ages so that the population pyramid is reversed and is characterised by a thin based but a broad top.

It may be noticed that indicators reflecting 'structure' of the population, 'state' of development and 'stage' of population transition are essentially very crude. This limitation is compelled by the very limited availability of data at the village level. More refined indicators of village population and development landscape could not be calculated simply because the necessary data are not available either through the decennial population census or from any other source. However, as the present analysis shows, even these crude indicators based on the limited data available at the village level can, quite effectively, depict the population and development landscape to study the population and development morphology in the villages of the country.

The five indicators reflecting the population and development landscape listed above have been estimated for each of the 517305 villages of the country on the basis of the data available through PCA 2011. The inter-village distribution of the five indicators is summarised in table 2 which shows that the distribution of different indicators across the villages of the country is essentially different. The coefficient of skewness is positive in all but one indicator which implies that in majority of the villages of the country, the value of these indicators is lower than the average. Similarly, in all the five indicators, the inter-village distribution is essentially platykurtic in shape as the value of excess kurtosis is negative. This means that the inter-village distribution of all the five indicators has broad peak but thin tails or the villages are fairly scattered away from the central tendency.

Table 2: Inter-village distribution of population and development indicators in India, 2011

Particulars	Illiteracy rate (*ILT*)	Proportion of non-workers (*NWR*)	Proportion of marginal workers (*MAR*)	Age structure index (*ASI*)	Fertility index (*FTI*)
	Very low (<13.89) Low (13.89-27.79) Average (27.79-41.68) High (41.68-55.57) Very high (≥55.57)	**Very low (<31.36) Low (31.36-47.99) Average (47.99-64.63) High (64.63-81.26) Very high (≥81.26)**	**Very low (<20.00) Low (20.00-40.00) Average (40.00-60.00) High (60.00-80.00) Very high (≥80.00)**	**Very low (<7.63) Low (7.63-12.17) Average (12.17-16.72) High (16.72-21.26) Very high (≥21.26)**	**Very low (<18.31) Low (18.31-30.64) Average (30.64-42.98) High (42.98-55.31) Very high (≥55.31)**
	Proportionate (Per cent) distribution of villages				
Very low	7.04	2.20	42.45	0.15	1.08
Low	32.09	28.03	22.58	12.13	36.15
Average	38.08	39.51	17.99	45.82	43.97
High	17.74	29.49	9.95	33.52	16.69
Very high	5.05	0.77	7.03	8.38	2.11
N	517305	517305	517305	517305	517305
	Summary measures of inter-village distribution				
Minimum	0.000	0.147	0.000	0.031	0.060
Median	0.313	0.557	0.262	0.142	0.337
Maximum	0.695	0.979	1.000	0.258	0.676
IQR	0.176	0.147	0.423	0.055	0.157
Skewness	0.327	-0.116	0.704	0.268	0.465
Excess Kurtosis	-0.212	-0.623	-0.477	-0.396	-0.236

Source: Computed by the author from Primary Census Abstract, 2011.

It is logical to assume that the five indicators used to characterise the population and development landscape may be correlated in the context of variation across villages. It is therefore imperative that indicators depicting the population and development landscape are grouped into domains in such a way that the correlation between indicators within a same domain is very high whereas the correlation between indicators belonging to different domains is the lowest. We have used the exploratory factor analysis procedure (Beavers et al, 2013; Sharma, 1996) for combining the five indicators of population and development landscape into mutually independent domains for constructing the population and development index. All the five indicators were first normalised to range between 0 and 1 and then the exploratory factor analysis procedure was applied on the basis of data from 517305 villages.

Results of the exploratory factor analysis procedure are presented in table 3. The exploratory factor analysis revealed that the five indicators reflecting the population and development landscape may be grouped into two distinct domains that account for almost 72 per cent of the total variance in the original data (Table 3). The KMO measure of sampling adequacy was greater than 0.60 which along with the Bartlett's test for sphericity suggest that the application of the exploratory factor analysis procedure was appropriate for combining the indicators and validates our approach of combing the indicators on the basis of the variation of the indicators across the villages of the country.

Table 3 indicates that the five indicators of population and development can be combined into two domains or factors representing the population and development landscape. The first factor or domain has high loadings in three indicators - *ASI*, *FTI* and *ILT* - and accounts for nearly 49 per cent of the total variance in the original data set or almost 68 per cent of the variance explained by the exploratory factor analysis model. This domain may be termed as the population scenario domain of the population and development landscape. The second domain, on the other hand, has high loadings in *NWR* and *MAR* and accounts for about 23 per cent of the total variance in the original data set or around 32 per cent of the variance explained by the exploratory factor analysis model. This domain may be termed as the development domain of the population and development landscape. The exploratory factor analysis thus suggested that the village population and development landscape can be characterised on a two-dimensional space - one reflecting the population structure and population transition while the other reflecting the state of development and the two dimensions of the population and development landscape are orthogonal to each other.

The exploratory factors analysis provides a deeper understanding of the population and development landscape at the village level. For example, table 3 suggests that the population scenario in the village is strongly associated with the level of education in the village but the level of education in the village has little impact on the level of participation of the village people in the productive activities. This means that the two components of development - education and work – does not appear to be related at the village level. One possible reason for this disconnect between work and education at the village level may be the suitability of the work for educated people. Data available through the 2011 population census suggest that almost three-fourth of the workers, either main or marginal, in the 517305 villages of the country included in the present analysis are either cultivators or agricultural labourers and working in the farm may not be an attractive preposition for the educated people.

Table 3 also indicates that the proportion of non-workers in the population (*NWR*) is negatively associated with the proportion of marginal workers (*MAR*) in the work force. This is expected as *NWR* and *MAR* reflect two essentially different aspects of participation in productive activities – the aspect of non-participation and the aspect of limited participation or casual participation respectively. Non-participation in productive activities implies that work is not available whereas participation for a limited period means that only casual work is available. The sum of the two proportions reflects the opportunity of participation in productive activities, the higher the sum, the lower the opportunity of full participation. The opportunity of full participation in productive activities may be low if the proportion of non-workers in the population is high but the proportion of marginal workers in the work force is low or if the proportion of non-workers in the population is low but the proportion of marginal workers within the work force is high irrespective of the negative correlation between them.

Table 3: Results of exploratory factor analysis

Factor or Domain I		Factor or Domain II	
Indicator	**Loadings**	**Indicator**	**Loadings**
Age structure index (*ASI*)	0.968	Proportion of non-workers (*NWR*)	0.838
Fertility index (*FTI*)	0.965	Proportion of marginal workers to total workers (*MAR*)	-0.574
Illiteracy rate (*ILT*)	0.671		
Sum of squared loadings	2.462	1.156	
Proportion of total variance explained	48.52	23.20	
KMO measure of sampling adequacy			0.606
Bartlett's test for sphericity			2144956

Source: Author's calculations based on Primary Census Abstract, 2011
Remark: Only those indicators are shown in each factor which has a factor loading of more than 0.55.

Results of the exploratory factor analysis suggest that an index reflecting the population domain of the population and development landscape or the population transition index (*PTI*) may be constructed as the weighted average of the normalised values of *ASI*, *FTI* and *ILT*. Similarly, an index that reflects the development domain of the population and development landscape or the development status index (*DSI*) may be constructed as the weighted average of the normalised values of *NWR* and *MAR*. In other words,

$$PTI = \omega_a a + \omega_f f + \omega_l l$$

where *a* is the normalised value of *ASI*, *f* is the normalised value of *FTI* and *l* is the normalised value of *ILT* and ω_a,ω_f and ω_l are weights to be estimated and $\omega_a+\omega_f+\omega_l=1$. Similarly,

$$DSI = \omega_n n + \omega_m m$$

where *n* and *m* are normalised values of *NWR* and *MAR* respectively and ω_n and ω_m are weights such that $\omega_n+\omega_m=1$. Finally, the composite population and development index (*PDI*) may be obtained as

$$PDI = \omega_T PTI + \omega_S DSI$$

where ω_T and $\omega_S$are weights to be estimated and $\omega_T+\omega_S=1$.

It is obvious that *PDI* as well as *PTI* and *DSI* range from 0 to 1. A high value of *PTI* reflects a young population with high fertility; and high illiteracy reflecting poor population scenario and an early stage of transition and vice versa. Similarly, a high value of *DSI* reflects low participation in productive activities and a high proportion of marginal workers in the work force which means poor state of development in the village and vice versa. Finally, the higher is the *PDI*, the poorer is the population and development landscape and vice versa.

It now remains to estimate weights required to calculate *PTI*, *DSI* and *PDI*. We follow the approach suggested by OECD (2008). Let x_a, x_f, x_l, x_n and x_m are loadings in factor 1 and y_a, y_f, y_l, y_n and y_m are loadings in factor 2 for *ASI*, *FTI*, *ILT*, *NWR* and *MAR* respectively. Also let v_1 is the variance explained by the factor 1 and v_2 is the variance explained by the factor 2. Then, different weights required for estimating *PTI*, *DSI* and *PDI* have been calculated as follows:

$$\omega_a = \frac{\left[x_a^2 * \left(\frac{v_1}{x_a^2+x_f^2+x_l^2}\right)\right]}{v_1}$$

$$\omega_f = \left[x_f^2 * \left(\frac{v_1}{x_a^2 + x_f^2 + x_l^2}\right)\right] / v_1$$

$$\omega_l = \left[x_l^2 * \left(\frac{v_1}{x_a^2 + x_f^2 + x_l^2}\right)\right] / v_1$$

$$\omega_n = \left[y_n^2 * \left(\frac{v_2}{y_n^2 + y_m^2}\right)\right] / v_2$$

$$\omega_m = \left[y_m^2 * \left(\frac{v_2}{y_n^2 + y_m^2}\right)\right] / v_2$$

$$\omega_T = \frac{v_1}{v_1 + v_2}$$

$$\omega_S = \frac{v_2}{v_1 + v_2}$$

Population and Development Landscape

Table 4 summarises the distribution of *PDI*, *PTI* and *DSI* across 517305 villages of the country. The inter-village distribution of all the three indexes is found to be positively skewed. This implies that the number of villages having the index *PDI* below the average is more than the number of villages having the index *PDI* above the average. Table 4 also suggests that in majority of the villages of the country, the state of development is relatively poor while the pace of population transition is relatively slow, although the degree of skewness is marginally higher in case of population transition than in case of the state of development. At the same time, the inter-village distribution of the *PDI* as well as *PTI* and *DSI* is platykurtic in shape having thin tails and broad peak. Moreover, the degree of the flatness in the inter-village distribution of *PTI* is substantially higher than the degree of flatness in the inter-village distribution of *DSI* as reflected by the higher value of the excess kurtosis. Table 4 also indicates that the inter-village distribution of the two dimensions of the population and development landscape is essentially different. This implies that there is no single indicator or dimension, either population or development, which can be used to characterise the village level population and development landscape. This observation justifies the use of the composite index to characterise the population and development landscape.

The village population and development landscape can be categorised into five categories on the basis of *PDI*: very good if $PDI<0.251$; good if $0.251 \leq PDI<0.418$; average if $0.418 \leq PDI<0.586$; poor if $0.586 \leq PDI<0.753$; and very poor if $PDI \geq 0.753$. This categorisation suggests that the population and development landscape appears to be poor or very poor in at least one fifth villages of the country whereas in more than one third villages, the population and development landscape may be categorised as good or very good. There are, however, a small proportion of villages in the country where the population and development landscape is either very poor or very good and, in majority of the villages, the population and development landscape may be termed as average.

Similarly, the population scenario in a village can be categorised on the basis of the index*PTI* as very good if $PTI<0.213$; good if $0.213 \leq PTI<0.408$; average if $0.408 \leq PTI<0.603$; poor if $0.603 \leq PTI<0.797$; and very poor if $PTI \geq 0.797$. According to this categorisation, the population scenario appears to be either very good or good in around 35 per cent of the villages; poor or very poor in more than 22 per cent villages and average in around 41 per cent villages. On the other hand, the state of development in a village may be categorised on the basis of the index*DSI* as very good if $DSI<0.278$; good if $0.278 \leq DSI<0.455$; average if $0.455 \leq DSI<0.631$; poor if $0.631 \leq DSI<0.808$; and very poor if $DSI \geq 0.808$. According to this categorisation, the state of development appears to be either very good or good in almost 40 per cent of the villages but poor or very poor in only about 10 per cent of the villages whereas in more than half of the villages of the country, the state of development, as measured through the index *DSI* may be categorised as average. However, there are very few villages where either the population scenario or the state of development is either very good or very poor.

Table 4: Inter-village distribution of *PDI*, *DTI* and *DSI*

PDI		***PTI***		***DSI***	
Very good (< 0.251)	1.20	Very good (< 0.213)	2.42	Very good (< 0.278)	3.28
Good (0.251-0.418)	33.09	Good (0.213-0.408)	34.01	Good (0.278-0.455)	36.36
Average (0.418-0.586)	44.85	Average (0.408-0.603)	41.01	Average (0.455-0.631)	51.46
Poor (0.586-0.753)	19.22	Poor (0.603-0.797)	19.03	Poor (0.631-0.808)	9.50
Very poor (≥0.753)	1.65	Very poor (≥0.797)	3.53	Very poor (≥0.808)	0.40
N	517305		517305		517305
Minimum	0.083		0.018		0.102
Median	0.472		0.466		0.487
Maximum	0.920		0.992		0.985
IQR	0.180		0.230		0.154
Skewness	0.298		0.367		0.140
Kurtosis	-0.441		-0.368		-0.073

Source: Author's calculations based on Primary Census Abstract, 2011

Population and Development Morphology

Is the population and development landscape of a village is influenced by the defining characteristics of the village? We have used the classification modelling approach (Tan, Steinbach, Kumar, 2006; Han, Kamber, Pei, 2012) to examine how *PDI* is related to selected village characteristics. Classification modelling involves classifying villages on the basis of *PDI* as the classification variable and selected village characteristics as predictor variables. The village characteristics used for classification included population size, gender balance and social class composition. The gender balance is measured in terms of male/female ratio while social class composition is measured in terms of the proportion of Scheduled Castes and proportion of Scheduled Tribes.

The classification and regression tree (CRT) method (Breiman et al, 1984) is used for classification modelling. CRT is a nonparametric method that divides villages into different categories so that within category homogeneity with respect to the classification variable is maximised. It recursively partitions the data space so that the partition can be represented as a decision tree (Loh, 2011). When the classification variable takes finite number of unordered values, the method generates classification tree. When the classification variable is either a continuous variable or an ordered discrete variable, regression tree is generated. The villages are sorted according to the classification variable into mutually exclusive categories based on the predictor variable that causes the most effective split on the basis of a similarity measure. The process is repeated until either the perfect similarity is achieved or the stopping criterion is met (Ambalavanan et al, 2006; Lemon et al, 2003). A category in which all villages have the same value of the classification or the dependent variable is termed as "pure." If a category is not "pure", then the impurity within the category can be measured through a number of impurity measures. We have used the Gini coefficient of impurity in the present analysis. The Statistical Package for Social Sciences (SPSS) has been used for classifying villages. The classification variable *PDI* is a continuous variable so that the regression tree has been generated.

Results of the classification modelling exercise are presented in table 5 while the classification tree is depicted in figure 1. The most important predictor variable is the proportion of Scheduled Tribes (*PST*) followed by the ratio of males to females (*MFR*). Compared to *PST*, the relative importance of *MFR* is 52.8 per cent but the importance of the population size (*POP*) is only 16.4 per cent and that of proportion Scheduled Castes (*PSC*) is just 10.3 per cent. The classification modelling has resulted in 15 partitions out of which 8 are terminal nodes or partitions, further partitioning of which is not possible. This means that 517305 villages can be partitioned into 8 groups or clusters of villages, each having distinct population and development morphology as may be seen from table 6. The average *PDI* is the lowest in cluster 13 (0.431) having 99994 (19.3 per cent) villages which means that population and development landscape in villages of this cluster is relatively the best, on average. By contrast, the population and development landscape is relatively the poorest in cluster 9 having 59374 (11.5 per cent) village and an average *PDI* of 0.537.

Table 6 presents the defining characteristics of villages in different clusters along with the population scenario and the state of development. The population and development landscape is relatively the best, on average, in villages of cluster 13. The defining characteristics of villages of this cluster include a gender balance favourable to females (*MFR*≤1.070) and a low proportion of the Scheduled Tribes population (0.001<*PST*≤0.267). On the other hand, the defining characteristics of villages of cluster 9, the cluster with highest *PDI*, include gender balance unfavourable to females (*MFR*>1.070), large population (*POP*>1228) and absence of Scheduled Tribes (*PST*≤0.001). *PDI* has also been found to be very high, in villages of cluster 4. The defining characteristics of villages of this cluster include a gender balance favourable to females (*MFR*≤1.070) but very high proportion of Scheduled Tribes (*PST*>0.576). Similarly, the second lowest *PDI*, on average, is found to be in villages of cluster 11 which are characterised by a gender balance highly favourable to females (*MFR*≤1.020) and the absence of Scheduled Tribes (*PST*≤0.001). It is obvious from table 6 that the defining characteristics of the village do impact upon the population and development landscape of the village.

Table 5: The classification table

Node	Village characteristics			*DPI*		N	Remarks
	Scheduled Tribes *PST*	Male/Female ratio *MFR*	Population size *POP*	Mean	Standard deviation		
0	*All*	*All*	*All*	0.481	0.123	517305	
1	*All*	≤1.070	*All*	0.466	0.122	309314	
2	*All*	>1.070	*All*	0.502	0.120	207991	
3	≤ 0.575	≤1.070	*All*	0.454	0.119	252683	
4	>0.575	≤1.070	*All*	0.523	0.121	56631	Terminal node
5	*All*	>1.070	≤1228	0.489	0.123	118130	Terminal node
6	*All*	>1.070	>1228	0.519	0.113	89861	
7	≤0.001	≤1.070	*All*	0.469	0.127	126705	
8	0.001-0.575	≤1.070	*All*	0.439	0.108	126478	
9	≤0.001	>1.070	>1228	0.537	0.109	59374	Terminal node
10	>0.001	>1.070	>1228	0.484	0.113	30487	Terminal node
11	≤0.001	≤1.021	*All*	0.456	0.127	73532	Terminal node
12	≤0.001	>1.021	*All*	0.487	0.125	52673	Terminal node
13	0.001-0.267	≤1.070	*All*	0.431	0.107	99994	Terminal node
14	0.267-0.575	≤1.070	*All*	0.468	0.106	26484	Terminal node

Source: Author’s calculations based on Primary Census Abstract, 2011

Table 6: Population and development morphology of villages in India

Particulars	Node (Village cluster)								All
	4	**5**	**9**	**10**	**11**	**12**	**13**	**14**	
	Cluster Characteristics								
PST	*>0.575*	*All*	*≤0.001*	*> 0.001*	*≤0.001*	*≤0.001*	*0.001-0.267*	*0.001-0.267*	*All*
MFR	*≤1.070*	*>1.070*	*>1.070*	*>1.070*	*≤1.021*	*>1.021*	*≤1.070*	*≤1.070*	*All*
POP	*All*	*≤1228*	*>1228*	*>1228*	*All*	*All*	*All*	*All*	*All*
	Population and Development Index (*PDI*)								
	Proportionate distribution of villages (per cent)								
Very good (< 0.251)	0.63	1.37	0.06	0.19	2.69	1.21	1.35	0.59	1.20
Good (0.251-0.418)	19.93	29.48	15.27	32.51	40.70	31.87	50.34	34.02	33.09
Average (0.418-0.586)	48.35	46.60	50.82	47.51	39.73	43.95	39.00	51.25	44.85
Poor (0.586-0.753)	28.15	20.65	31.60	18.58	15.50	20.97	8.69	13.25	19.22
Very poor (≥0.753)	2.93	1.89	2.26	1.21	1.37	2.00	0.62	0.90	1.65
	Summary measures of inter-village distribution								
Minimum	0.116	0.109	0.179	0.098	0.083	0.101	0.110	0.153	0.083
Median	0.520	0.483	0.538	0.474	0.443	0.479	0.413	0.459	0.472
Maximum	0.920	0.912	0.896	0.881	0.892	0.902	0.917	0.875	0.920
IQR	0.174	0.177	0.159	0.167	0.187	0.186	0.144	0.144	0.180
Skewness	0.078	0.221	0.065	0.412	0.388	0.260	0.702	0.454	0.298
Kurtosis	-0.437	-0.412	-0.516	-0.407	-0.395	-0.512	0.310	0.042	-0.441
	Population Transition Index (*PTI*)								
	Proportionate distribution of villages (per cent)								
Very good (< 0.213)	0.92	2.73	0.29	0.72	5.13	2.67	2.94	1.10	2.42
Good (0.213-0.408)	18.52	30.75	19.25	31.65	42.44	34.31	49.52	34.79	34.01
Average (0.408-0.603)	42.84	41.36	46.35	45.50	36.51	39.79	37.35	47.13	41.01
Poor (0.603-0.797)	29.66	21.01	29.63	19.39	13.64	19.56	9.01	14.97	19.03
Very poor (≥0.797)	8.05	4.15	4.48	2.74	2.28	3.66	1.18	2.00	3.53
	Summary measures of inter-village distribution								
Minimum	0.030	0.018	0.060	0.041	0.028	0.041	0.033	0.036	0.018
Median	0.548	0.480	0.538	0.474	0.418	0.465	0.399	0.456	0.465
Maximum	0.991	0.992	0.962	0.961	0.986	0.979	0.976	0.966	0.992
IQR	0.233	0.235	0.212	0.206	0.227	0.238	0.182	0.187	0.230
Skewness	0.122	0.262	0.144	0.429	0.514	0.348	0.708	0.500	0.367
Kurtosis	-0.480	-0.452	-0.529	-0.291	-0.167	-0.443	0.422	0.038	-0.368
	Development Status Index (*DSI*)								
	Proportionate distribution of villages (per cent)								
Very good (< 0.278)	5.71	3.97	0.30	1.26	3.96	2.60	3.27	3.58	3.28
Good (0.278-0.455)	43.03	33.03	15.63	41.63	32.17	27.48	48.05	42.92	35.36
Average (0.455-0.631)	44.15	52.12	74.13	50.23	49.04	57.44	41.55	46.96	51.46
Poor (0.631-0.808)	6.85	10.30	9.82	6.81	13.96	11.98	6.92	6.40	9.50

Very poor (≥0.808)	0.26	0.58	0.11	0.07	0.88	0.50	0.21	0.14	0.40
	Summary measures of inter-village distribution								
Minimum	0.102	0.102	0.155	0.179	0.102	0.104	0.116	0.115	0.102
Median	0.459	0.488	0.526	0.473	0.501	0.507	0.451	0.465	0.487
Maximum	0.939	0.985	0.882	0.871	0.973	0.975	0.930	0.936	0.985
IQR	0.170	0.147	0.094	0.140	0.179	0.139	0.153	0.156	0.154
Skewness	0.173	0.167	-0.039	0.241	0.106	0.013	0.434	0.179	0.140
Kurtosis	-0.232	0.081	0.817	-0.208	-0.312	0.073	-0.003	-0.245	-0.073
N	56631	118130	59374	30487	73532	52673	99994	26484	517305

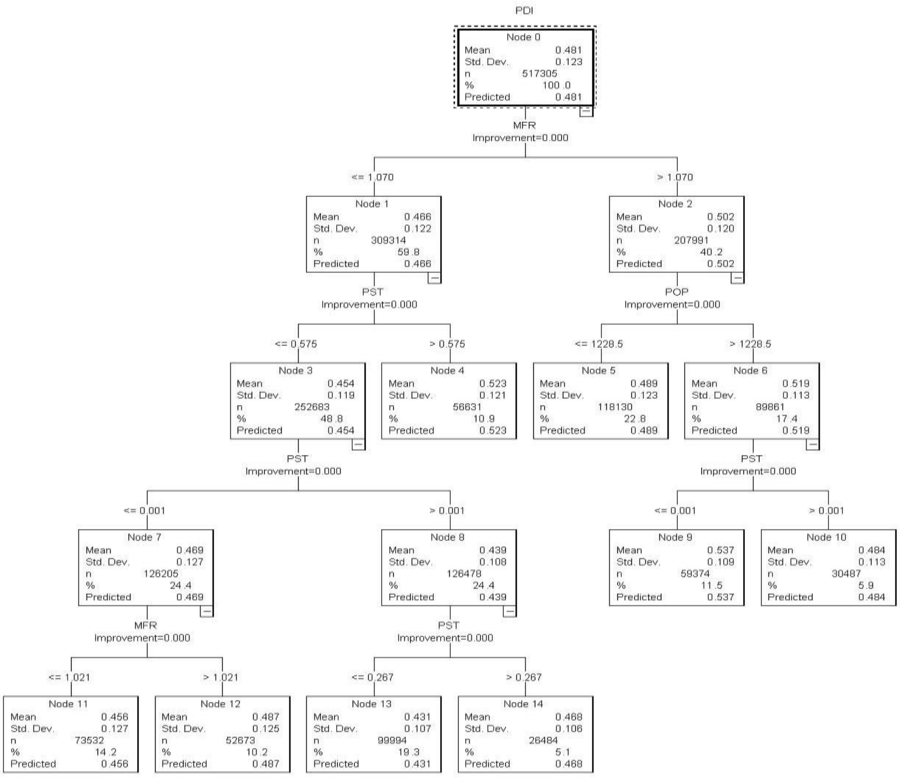
Source: Author’s calculations based on Primary Census Abstract, 2011

Table 7: Distribution of villages in states/Union Territories by population and development morphology

Country/State	Proportion (per cent) of villages in node (cluster)								Number of villages
	4	5	9	10	11	12	13	14	
Jammu & Kashmir	5.41	28.57	5.33	14.00	16.09	12.59	14.00	4.02	5401
Himachal Pradesh	2.74	26.11	0.68	1.00	41.02	12.33	14.02	2.09	11742
Punjab	0.00	33.11	31.58	0.00	16.53	18.78	0.00	0.00	11443
Uttarakhand	1.65	20.97	4.72	2.23	50.12	10.44	8.11	1.76	8254
Haryana	0.00	23.80	62.20	0.00	3.03	10.97	0.00	0.00	6209
Delhi	0.00	12.37	83.51	0.00	1.03	3.09	0.00	0.00	97
Rajasthan	9.49	32.20	7.05	10.82	8.82	8.68	18.82	4.13	37919
Uttar Pradesh	0.16	27.58	29.73	2.83	19.82	14.08	5.69	0.12	88207
Bihar	0.79	21.67	31.73	7.65	11.54	15.66	10.42	0.53	35099
Sikkim	12.40	33.77	0.00	11.08	0.26	0.00	19.26	23.22	379
Arunachal Pradesh	60.94	29.47	0.05	1.25	1.60	0.45	2.90	3.35	2002
Nagaland	61.79	25.44	0.00	10.37	0.00	0.00	0.35	2.04	1128
Manipur	54.77	22.64	0.36	2.94	9.80	2.58	6.50	0.41	1939
Mizoram	58.74	36.71	0.00	4.02	0.00	0.00	0.35	0.17	572
Tripura	38.08	5.51	1.44	17.60	1.32	2.75	23.23	10.06	835
Meghalaya	67.57	28.53	0.10	0.43	0.91	0.46	1.04	0.96	3950
Assam	15.70	21.60	5.20	3.41	15.33	13.78	19.53	5.46	21778
West Bengal	6.14	20.38	6.97	6.92	10.06	14.00	28.04	7.50	35224
Jharkhand	29.42	25.28	3.35	5.07	7.54	7.10	12.17	10.07	24558
Odisha	20.68	20.56	1.95	2.77	13.43	7.83	20.16	12.62	37526
Chhattisgarh	33.18	10.52	0.16	2.27	5.05	2.29	29.21	17.32	17459
Madhya Pradesh	18.64	32.86	3.90	9.66	4.89	4.61	17.40	8.05	45659
Gujarat	19.98	15.47	9.10	9.61	10.12	14.23	16.70	4.78	17013
Maharashtra	11.10	20.53	3.41	14.03	8.80	5.37	30.62	6.14	38117
Andhra Pradesh	9.48	7.15	1.05	4.07	14.92	6.54	51.46	5.32	23232
Karnataka	2.20	14.82	1.16	4.70	17.61	7.90	45.84	5.76	25229
Goa	5.82	12.33	1.71	6.51	17.12	8.90	35.27	12.33	292
Kerala	0.20	0.00	0.00	0.40	4.65	0.10	92.87	1.78	1010
Tamil Nadu	1.22	6.08	4.46	1.89	43.75	16.36	25.82	0.41	14630
Union Territories	15.92	26.87	3.48	8.71	28.36	7.71	7.71	1.24	402
India	10.95	22.84	11.48	5.89	14.21	10.18	19.33	5.12	517305

Source: Author’s calculations based on Primary Census Abstract, 2011

Figure 1: Classification of villages according to *PDI* on the basis of the defining characteristics of the village.



It is very much evident from table 6 that the defining characteristics of the village influence both the population scenario and the state of development of the village and hence the population and development landscape. It appears that the gender balance and the proportion of Scheduled Tribes in the village have substantial impact on the population and development landscape of the village. In villages where the gender balance is relatively favourable to females, the population and development landscape is definitely better than the population and development landscape in those villages where the gender balance is unfavourable. Table 6 also indicates that the population and development landscape is the poorest in those villages which are large in population size and where the gender balance is unfavourable to females. Similarly, the population and development landscape is very poor in villages with a high concentration of Scheduled Tribes population even if the gender balance in these villages is relatively favourable to females.

Any discussion on population and development morphology in India is incomplete without the regional analysis. Table 7 presents state/Union Territory wise distribution of villages across different clusters to reflect the population and development morphology in different states/Union Territories of the country. In many states/Union Territories, there is very high concentration of villages in particular clusters. For example, almost 93 per cent villages in Kerala belong to cluster 13, the cluster which has the lowest *PDI* on average meaning relatively the best population and development landscape. In Andhra Pradesh and Karnataka more than 40 per cent villages belong to this cluster whereas in West Bengal, Odisha, Chhattisgarh, Maharashtra, Goa and Tamil Nadu a substantial proportion of villages belong to this cluster. On the other hand, more than 60 per cent villages in Haryana and Delhi belong to cluster 9 whereas, at least 40 per cent of villages in the north-eastern states of Arunachal Pradesh, Nagaland, Manipur, Mizoram, and Meghalaya belong to cluster 4. The *PDI*, on average, is the highest and the second highest in these clusters reflecting poor to very poor population and development landscape. Similarly, at least 40 per cent villages in Himachal Pradesh, Uttarakhand and Tamil Nadu belong to cluster 11 where *PDI* is quite low, on average, and reflects relatively better population and development landscape. On the other hand, in many states, villages are almost equally distributed across more than two clusters reflecting contrasting population and development landscape. For example, more than 60 per cent villages in Punjab are distributed almost equally between cluster 5 and cluster 9. The population and development index *PDI* is relatively the lowest, on average, in cluster 9 whereas it is quite low, on average, in cluster 5. This shows that there is considerable inequality in population and development landscape across villages in Punjab. A similar situation prevails in other states also including Uttar Pradesh and Bihar.

Discussions and Conclusions

The present analysis demonstrates that village population and development landscape in India is conditioned by the defining characteristics of the village such as population size, gender balance and social class composition. The characteristics of the village are essentially exogenous to the village social and economic development system. They are deep rooted in the culture and tradition of the Indian society. This means that the cultural and traditional divisions of the Indian society remain largely unaffected by population and development efforts in the rural India. The dividends of population and development efforts appear to be confined to only those villages which are large in size, where gender balance is favourable to females and where Scheduled Tribes are not in majority. There are however very few villages which meet all the three conditions. More than half of the villages in India are small with a population of less than 1000. Similarly, the gender balance is unfavourable to females in more than 60 per cent villages. There appears little impact of population and development efforts in these villages as the population and development landscape in these villages appears to remain relatively poor to very poor.

The social class effects of population and development appear to be even more dominating. The population and development landscape is found to be very poor in villages with a heavy concentration of Scheduled Tribes irrespective of the fact that the gender balance in these villages is favourable to females. The population and development landscape has been found to be substantially different in villages where Scheduled Tribes constitute more than 60 per cent of the population as compared to villages where the proportion of Scheduled Tribes is less than 60 per cent. In fact, the population and development landscape is found to be relatively the best in villages where Scheduled Tribes constitute just around one fourth of the village population.

The analysis also confirms that education matters as far as population transition is concerned. This means that universalisation of education in Indian villages may contribute significantly towards hastening the pace of population transition. However, a more revealing finding of the present analysis is that neither the level of education nor the stage of population transition appears to have any impact on the participation of the people in village level social and economic productive activities. It appears that opportunities of participation in social and economic productive activities are limited at the village level, especially for an educated person. The village economy in India remains primarily agrarian and traditional. Workers engaged in agriculture related activities account for almost three fourth of the rural work force in the country. A transformation of the village social and economic production system is widely recognised as a prerequisite for increasing participation, especially of females, in the village level productive activities. There is however little indication of such a transformation in rural India. As a result, whatever limited demographic bonus is generated from population transition, it largely remains un-reaped in productive terms at the village level.

The present analysis also indicates that the regional variation in population and development morphology of villages is largely an offshoot of the regional variation in the defining characteristics of villages. This appears to be the reason why population and development efforts have not been able to reduce regional population and development disparities that are so pervasive in India. The contrasting population and development landscape in villages of different characteristics suggests that villages having different defining characteristics require different approaches to addressing population and development issues.

From the policy perspective, the analysis calls for a spatial approach to population and development planning and programming with the village as the basic unit so that the defining characteristics of the village are duly taken into consideration at the planning stage. The first and perhaps the most important prerequisite for institutionalising the spatial approach to population and development planning is the decentralisation of the public and development administration system. This is challenging in India as, despite all talks of decentralisation at the policy level, the public and development administration system in the country remains the traditional command and control system introduced during the colonial period.

References

Ambalavanan, N, Carlo, WA, Shankaran, S, Ban, CM, Emrich, SL, Higgins, RD, Tyson, JE, O'Shea, TM, Laptook, AR, Ehrenkranz, RA, Donovan, EF, Waleh, MC, Goldberg, RN, Das, A, National Institute of Child Health and Human Development Neonatal Research Network (2006) 'Predicting outcomes of neonates diagnosed with hypoxemic-ischemic encephalopathy',*Pediatrics,118*(5), pp. 2084-93.

Bailey, FG (1957) *Caste and Economic Frontier*. Manchester, Manchester University Press.

Bailey, FG (1963) *Politics and Social Change: Orissa in 1959*. Berkeley, University of California Press.

Bauer, G, Davies, JA, Pelikan, J, Noack, H, Broesskamp, U, Hill, C (On behalf of the EUPHID Consortium) (2003) 'Advancing a theoretical model for public health and health promotion indicators development. Proposal from the EUPHID Consortium',*European Journal of Public Health,13*(3), pp. 107-113.

Beavers, AS, Lounsbury, JW, Richards, JK, Huck, SW, Skolits, GJ, Esquivel, SL(2013) 'Practical considerations for using exploratory factor analysis in educational research',*Practical Assessment Research and Evaluation,18*(6). http://pareonline.net/getvn.asp?v=18&n=6.

Beteille, A (1965) *Caste, Class and Power: Changing Patterns of Stratification in a Tanjore Village*. Berkeley, University of California Press.

Beteille, A (1969) *Castes Old and New: Essays in Social Structure and Social Stratification*. Bombay, Asia Publishing House.

Bloom, DE, Canning, D, Sevilla, J, (2001) 'Economic growth and the demographic transition', National Bureau of Economic Research. Working Paper 8685.

Brieman, L, Friedman, JH, Olshen, RA, Stone, CJ, (1984) *Classification and Regression Trees*. CRC Press.

Campbell, G (1852) *Modern India, A Sketch of the System of Civil Government to which is Prefaced Some Account of the Natives and Native Institutions*. London, J. Murray.

Chakravarti, A (1975) *Contradiction and Change: Emerging Patterns of Authority in a Rajasthan Village*. Delhi: Oxford University Press.

Chaurasia, AR (2016) 'Child deprivation in Gujarat',*Social Change,46*(3), pp. 371-392.

Davis, K (1955) 'The origin and growth of urbanization in the world',*American Journal of SociologyIX*.

Doxiadis, CA (1968) *Ekistics: An Introduction to the Science of Human Settlements*. New York, Oxford University Press.

Durkheim, E (1982) *The Rules of Sociological Method*. New York, Free Press.

Gandhi, MK (1944) *Hind Swaraj*. Ahmedabad, Navjivan Press.

Ghosh, A, Singh, LL (2004) 'Child immunisation in rural West Bengal (India): A multilevel analysis', Paper presented in the Population Association of America, 2004 Annual Meeting, Boston.

Government of India (2011) *Census of India. Provisional Population Totals. Rural-Urban Distribution of Population. India, Series 1.* New Delhi, Office of the Registrar General and Census Commissioner. Paper 2, Volume 1 of 2011.

Government of India (2011a) *Census of India 2011. Metadata*. New Delhi, Office of the Registrar General and Census Commissioner.

Han, J, Kamber, M, Pei, J (2012) *Data Mining. Concepts and Techniques*. Amsterdam, Elsevier.

Hanafin, S, Brooks, AM (2005) *Report of the Development of a National Set of Child Well-being Indicators in Ireland*. Dublin, The National Children's Office.

Kessinger, T (1971) *Vilyatpur 1848-1968: Social and Economic Change in a North Indian Village*. Berkeley, University of California Press.

Kravdal, Ø (2004) 'Child mortality in India: the community-level effect of education',*Population Studies*,*58*(2), pp. 177-192.

Lemon, SC, Roy, J, Clark, MA, Friedman, PD, Rakowski, W (2003) 'Classification and regression tree analysis in public health: methodological review and comparison with logistic regression',*Annals of Behavioral Medicine*,*26*(3), pp. 172-81.

Loh, WY (2011) 'Classification and regression trees', *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*,*1*, pp. 14-23.

McNay, K, Arokiasamy, P, Cassen, RH (2003) 'Why uneducated women in India using Contraception? A Multilevel Analysis',*Population Studies*,*57*(1), pp. 21-40.

Nehru, JL (1961) *The Discovery of India*. Bombay, Asia Publishing House.

OECD (2008) *Handbook on Constructing Composite Indicators. Methodology and User Guide*. Paris, Organization for Economic Cooperation and Development.

Redfield, R (1956) *Peasant Society and Culture: An Anthropological Approach to Civilisation*. Chicago, University of Chicago Press.

Shah, AM (1973) *The Household Dimensions of the Family in India*. Berkeley, University of California Press.

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

Singh, A, Chakraborty, S, Roy, TK (2008) 'Village size and development in India', *Asian Population Studies*

Srinivas, MN (1952) *Religion and Society among the Coorgs of South India*. Bombay, Asia Publishing House.

Stephenson, R, Tsui, AO (2002) 'Contextual influences on reproductive health service use in Uttar Pradesh, India',*Studies in Family Planning*,*33*(4), pp. 309-320.

Tan, P-N, Steinbach, M, Kumar, V (2006) *Introduction to Data Mining*. New Delhi, Pearson Education.

Thakur, M (2014) 'Understanding ruralities: contemporary debates'. New Delhi, Jawaharlal Nehru University. Centre for Studies of Social Systems. Working Paper CAS/WP/14-1.