Journal of Geography and Regional Planning Vol. 5(12), pp. 314-327, 30 June, 2012
Available online at http://www.academicjournals.org/JGRP
DOI: 10.5897/JGRP11.078
ISSN 2070-1845 ©2012 Academic Journals

Full Length Research Paper

Demographic diversity in India: Evidence from the provisional results of 2011 population census

Aalok Ranjan Chaurasia

Shyam Institute, 82 Aradhana Nagar, Bhopal, MP-462003, India.
E-mail: aranjan@shyaminstitute.in.

Accepted 26 January, 2012

The present paper uses provisional figures available through the 2011 population census to analyse the demographic diversity across the districts of India. The demographic diversity is measured in terms of a diversity index that has been developed for the purpose. The analysis suggests that most of the demographic diversity across the districts of the country is confined to a few districts only, despite the wide spread social, cultural, economic and ecological diversity that is so pervasive in India. The analysis also suggests that between States, demographic diversity accounts for a larger proportion of the total inter-district diversity in the country as compared to within States inter-district diversity and there are substantive inter-State variations. Demographic diversity across the States/Union Territories of India is well known. The provisional results of the 2011 population census suggest that this diversity continues to persist.

Key words: India, population, census, demographic, ecological diversity.

INTRODUCTION

India is a very diverse country demographically as well as in terms of social and economic development. There are variations not only across States and Union Territories but also across districts within States and Union Territories and sub-districts within districts, etc. In view of this wide ranging diversity across the country, there is a renewed emphasis in recent years to promote decentralised districts-based approach to population and development planning. This emphasis is amply reflected in the 73rd and 74th amendments in the Constitution of India, 10th and 11th five-year development plans (Government of India, 2002, 2008), National Population Policy 2000 (Government of India, 2000), National Health Policy 2002 (Government of India, 2002) and the National Rural Health Mission (Government of India, 2005).

The diversity in demography and development across the country and the emphasis on decentralised district population and development planning calls for analysing how the demographic and development situation varies across the districts and how the demographic and development situation in a district contributes to the situation at State/Union Territory level and at the national level. Analysing the demographic and development diversity is thus important through a policy perspective as the persistence of the diversity suggests that exogenous variables and policy and programme interventions affect the demographic and development situation differently at the lower level administrative units - districts, sub-districts and even towns and villages.

In this paper, we use the provisional figures of the 2011 population census to analyse demographic diversity across the districts of India. The provisional results of the 2011 population census provide information about population and its key characteristics for 640 districts of India as they existed at the time of the census. Although, carried out at an interval of ten years, the decennial population census is the only source of district level demographic information in India to facilitate assessment of the demographic situation. The population census is also the only source to provide population related information to the decentralised district population and development planning process and for evaluating the impact of population and development programmes and activities on the quality of life of the people at the district and below district levels.

This paper is organized as follows. In the next section of the paper, we develop a diversity index to measure the demographic diversity and show how the demographic diversity at the district level contributes to the demographic diversity at the State/Union Territory level and at the national level. Estimates of the diversity index for the country, States/Union Territories and districts are presented in section three of the paper while section four analyses the variation in the index in the context of selected demographic variables has been analysed in section four. Finally, the policy and programme implications of the diversity in the demographic situation across the districts of the country in the context of decentralised district population and development planning have been discussed in the fifth section of the paper.

DEMOGRAPHIC DIVERSITY INDEX

We measure the demographic diversity across the districts on a two dimensional scale, the dimension of the extent or intensity of diversity and the dimension of the extensiveness of diversity. Measures of the extent or intensity of diversity include differentials and concentration. Differentials are the most basic. They measure how a demographic variable V in the district d, V_d, deviates from the national average, V_c. If $V_d/V_c=1$ for all districts in the country, there is no diversity in the variable V across districts. The larger is the deviation from the limiting value of 1 and the larger is the disparity across the districts. The ratio V_d/V_c, therefore, is an indicator of the extent or the intensity of inter-district diversity of the variable V in relation to the situation at the national level.

One problem in using the ratio V_d/V_c as an indicator of the extent or intensity of diversity is that it may take exorbitantly high or low values for extremely high or low values of V_d. This concern can be circumvented by using the logarithmic scale rather than the normal scale. Thus, the index of intensity of a demographic variable V in district d relative to the intensity in the country as whole may be defined as:

$$I_{dc(v)} = \log(V_d/V_c) \qquad (1)$$

where log stands for logarithm to the base 10, so that log(1)=0. It is obvious that $I_{dc(v)}=0$ when $V_d/V_c=1$; $I_{dc(v)}>0$ when $V_d/V_c>1$; and $I_{dc(v)}<0$ when $V_d/V_c<1$.

On the other hand, the extensiveness of the diversity may be measured simply in terms of the population in district d as proportion of the population of the country as a whole. If P_d denotes the population of the district d and P_c denotes the population of the country, then the extensiveness of the population in district d relative to the population of the country may be defined as:

$$E_{dc} = P_d/P_c \qquad (2)$$

It is obvious that $\sum E_{dc} = 1 \ \forall d \in c$.

The dimension of extensiveness in the measurement and analysis of diversity is important because the population and the geographical area are not the same for all districts and this structural diversity may influence the demographic diversity. Accounting for the structural diversity is therefore necessary for any analysis of demographic diversity across districts.

Using the index of the extensiveness of diversity and the index of the extent or intensity of the diversity of variable V in district d, we define the index of diversity in district d in relation to the situation in the country as a whole as:

$$D_{dc(v)} = E_{dc} \times (I_{dc(v)})^2 \qquad (3)$$

It may be noticed that the index $D_{dc(v)}$ is always positive. The larger the value of the index $D_{dc(v)}$, the higher is the diversity in the district d as compared to the situation at the national level. The index $D_{dc(v)}$ is a fuller measure of diversity in variable V across districts as it takes into account the size of the population of the district.

Finally, total diversity in variable V across all districts of the country may now be defined as:

$$D_{cd(v)} = \sum D_{dc(v)} = \sum E_{dc}*(I_{dc(v)})^2 \quad \forall\ d \in c \qquad (4)$$

which is nothing but the weighted sum of the square of the index of intensity of the variable V in the districts of the country. It is clear that the index $D_{cd(v)}$ takes into account the two dimensions of diversity, the dimension of relative diversity and the dimension of extensiveness, as discussed earlier.

Arguing on similar lines, the total diversity in variable V in all districts of a State/Union Territory s within the country may also be defined as:

$$D_{sd(v)} = \sum D_{ds(v)} \quad \forall\ d \in s \qquad (5)$$

$$D_{ds(v)} = E_{ds}*(I_{ds(v)})^2 \qquad (6)$$

$$I_{ds(v)} = \log(V_d/V_s) \qquad (7)$$

$$E_{ds} = P_d/P_s \qquad (8)$$

Similarly, we can also define,

$$D_{cs(v)} = \sum D_{sc(v)} \quad \forall\ s \in c \qquad (9)$$

$$D_{sc(v)} = E_{sc}*(I_{sc(v)})^2 \qquad (10)$$

$$I_{sc(v)} = \log(V_s/V_c) \qquad (11)$$

$$E_{sc} = P_s/P_c \quad (12)$$

It is now easy to show that:

$$D_{dc(v)} = E_{sc}*D_{ds(v)} + E_{ds}*D_{sc(v)} + 2*E_{sc}*I_{ds(v)}*I_{sc(v)} \quad (13)$$

and

$$D_{dc(v)} = \sum E_{sc}*D_{ds(v)} + \sum E_{ds}*D_{sc(v)} + 2 \sum E_{sc}*I_{ds(v)}*I_{sc(v)} \quad (14)$$

Equation (14) decomposes the total diversity in variable V across the districts in the country into two components: diversity across districts within a State/Union Territory or within State/Union Territory component and diversity across States/Union Territory within the country or between State/Union Territory component, and an interaction term. The interaction term may be distributed across the within and between States/Union Territory components of the diversity following the Goldberg's rule (Durand, 1948).

Data and variables

The analysis presented here is based on the provisional figures of the 2011 population census released by the Registrar General and Census Commissioner of India (Government of India, 2011). These figures are related to the total population, population aged 0 to 6 years and population aged 7 years and above for each of the 640 districts of the country as well as for its 28 States and 7 Union Territories separately for males and females and for both sexes combined.

The provisional figures of the 2011 population census permit estimation of the following 7 demographic indicators for all districts, States and Union Territories of the country:

1. Population density- It is defined as the population per square kilometer of the geographical area. Population density is the most commonly used indicator of the distribution of population across administrative units. If there is no change in the administrative boundaries, the change in the population density of an administrative unit is proportional to the change in the population size of the administrative area. If the population of a district has increased by 10%, population density of that district would also increase by 10% if there is no change in the administrative boundary of the district.
2. Proportion of the population aged 0 to 6 years to the total population. This proportion is a crude indicator of the age structure of the population. The higher is the ratio, the younger is the age structure of the population.
3. The index of age composition. This indicator is defined as the ratio of the population aged 0 to 6 years to the population aged 7 years and above. The index of age composition may be regarded as a crude indicator of the age structure of the population.
4. Population sex ratio, It is defined as the ratio of females to males of all ages combined. This indicator reflects the sex balance in the population. In case of sex imbalance on the either side, the population sex ratio deviates from the limiting value of 1 when the number of females are equal to the number of males in the population.
5. Sex ratio in the population aged 0 to 6 years and sex ratio in the population aged 7 years and above. Age specific sex ratios normally differ from the population sex ratio. The sex ratio tends to be low at very young ages and increases with increasing age. 'Young' population and population with high birth rate tend to have lower overall sex ratio than 'Old' population and population with low birth rate (Shryock and Siegel, 1976).
6. Fertility index- It is defined as the ratio of the population aged 0 to 6 years to females aged 7 years and above. This ratio is similar to the conventional child-women ratio (Shryock and Siegel, 1976), although, it includes women which are not exposed to the risk of a birth. This ratio gives an idea about the diversity in fertility levels across administrative units.

It is well known that the demographic indicators described above are influenced by the core demographic processes: fertility, mortality and migration. As such, the underlying assumption of the present analysis is that diversity in the six variables across districts and States/Union Territories of the country broadly reflect the inter-district demographic diversity in the country. It may however be reiterated that the 2011 population census data released so far is only provisional. There may be changes in these data when final figures of the population census are released. It will however take some time for the Registrar General and Census Commissioner of India to release final data for all the 640 districts, 28 States and 7 Union Territories of the country. Till then, the only source of data for analysing demographic diversity in India is the provisional data of the 2011 population census.

DEMOGRAPHIC DIVERSITY IN INDIA 2011

Estimates of six demographic variables for India are given in Table 1. These estimates have been derived from the provisional figures of the 2011 population census. The table also includes summary measures of the inter-district variations in the six variables, whereas the kernel density plots of the distribution of the variables across the districts of the country are presented in Figure 1. These summary measures suggest that all the six demographic variables vary widely across the districts of the country. This is expected because of the social, cultural, economic and environmental diversity that is so pervasive in India. Moreover, the distribution of the six

Table 1. Selected demographic indicators and associated diversity index in India and summary measures of inter-district variations.

Indicator	Population density	Index of age composition	Sex ratio			Fertility index
			Population	Population aged 0 to 6 years	Population aged 7 years and above	
			Estimates			
India	381	0.151	0.940	0.914	0.944	0.311
			Summary measures of inter-district variations			
Minimum	1	0.072	0.533	0.774	0.500	0.152
Q1	210	0.124	0.905	0.892	0.904	0.254
Median	379	0.154	0.947	0.926	0.950	0.316
Q3	719	0.187	0.981	0.953	0.987	0.388
Maximum	45594	0.290	1.176	1.013	1.206	0.636
Range	45593	0.218	0.643	0.239	0.706	0.484
IQR	509	0.064	0.076	0.062	0.083	0.134
Mean	974	0.157	0.943	0.918	0.947	0.325
5% trimmed mean	486	0.156	0.943	0.920	0.947	0.322
Std. deviation	3529	0.041	0.062	0.043	0.068	0.086
Skewness	8.890	0.498	-0.558	-0.744	-0.537	0.480
Kurtosis	89.444	-0.263	5.125	0.004	5.422	-0.253
			Diversity index ($\times 10^{-5}$)			
India (D_{cd})	23526.4	1264.3	61.3	37.8	70.4	1383.3
			Summary measures of inter-district variations in the index (D_{dc})			
Minimum	0.000	0.000	0.000	0.000	0.000	0.000
Q1	1.904	0.154	0.007	0.005	0.007	0.131
Median	7.016	0.790	0.032	0.023	0.035	0.861
Q3	19.107	2.502	0.104	0.067	0.123	2.672
Maximum	2128.224	38.852	2.951	0.756	3.369	36.103
Range	2128.224	38.852	2.951	0.756	3.369	36.103
IQR	17.236	2.351	0.098	0.063	0.116	2.542
Mean	36.760	1.975	0.096	0.059	0.110	2.161
5% trimmed mean	14.976	1.497	0.063	0.044	0.072	1.654
Std. deviation	145.953	3.295	0.207	0.095	0.241	3.499
Skewness	8.841	4.507	6.716	3.269	6.625	3.652
Kurtosis	94.331	34.264	67.874	13.967	64.854	21.801
N	640	640	640	640	640	640

Source: Author's calculations.

demographic variables across the districts is essentially different as may be seen from the values of the skewness and kurtosis of these distributions. The distribution of population density across the districts is highly positively skewed and has a very high value of kurtosis. This suggests that some districts of the country have exceptionally high population density. At the same time, a very high value of kurtosis indicates that there is a very high probability of districts with extremely high population density. This observation is also supported by the kernel density plot in Figure 1. According to the provisional figures of the 2011 population census, the population density was more than 2000 persons per square kilometre in about 15% districts of the country, whereas in 11 districts, it was estimated to be more than 10 thousand persons per square kilometre. District Mumbai in Maharashtra had the highest population density of 50 thousand per square kilometre in the country, whereas in all the districts of the national capital Territory of Delhi, the population density was estimated to be more than 3500 persons per square kilometre.

The distribution of the index of age composition and

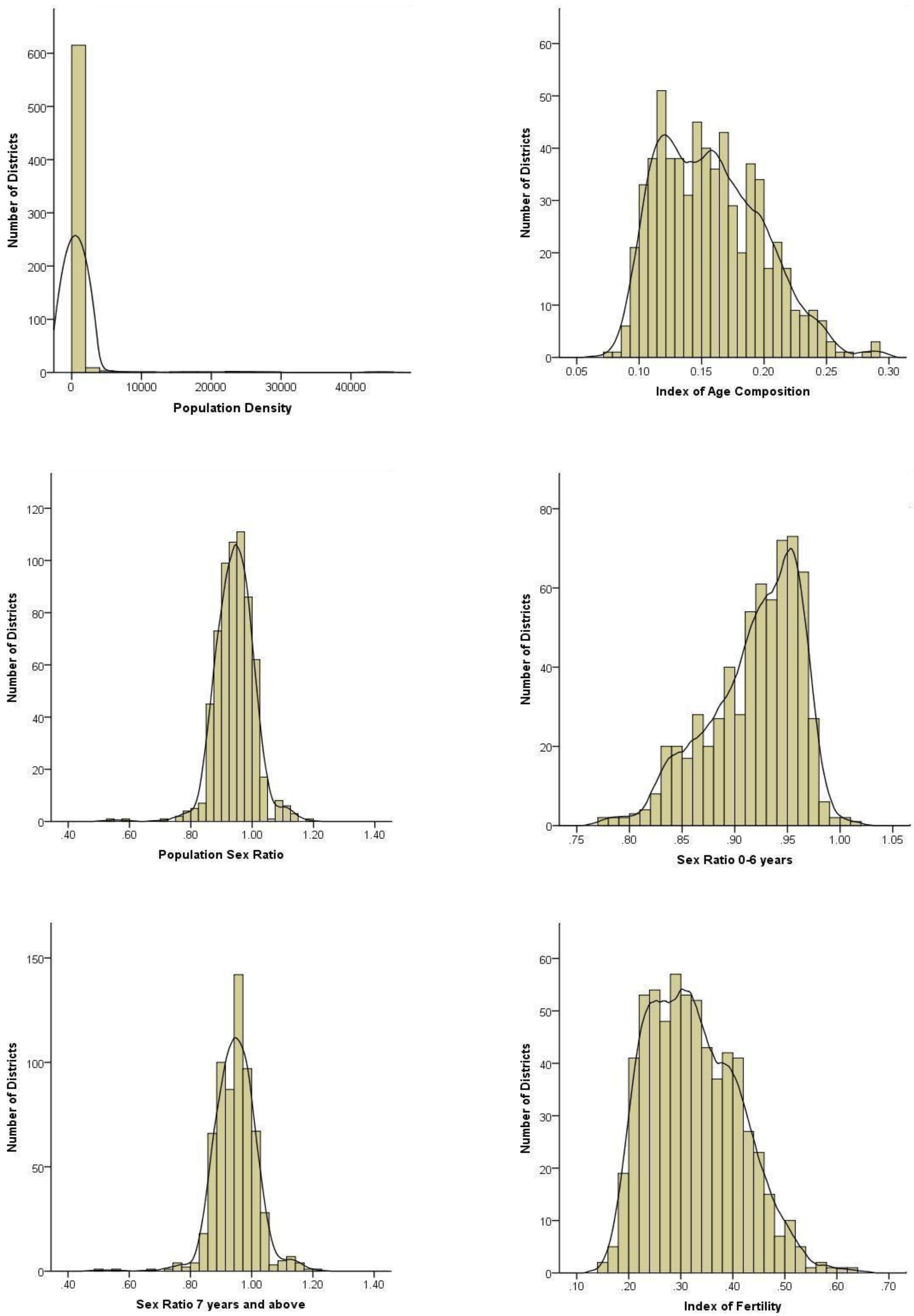

Figure 1. Kernel density plots of selected demographic variables in India, 2011.

the index of fertility across the districts of the country, on the other hand, appears to be very similar. The skewness is positive, but not very large in both distributions which suggests that there are only a few districts in the country with extremely high ratio of the population aged 0 to 6 years to the population aged 7 years and above. At the same time, the kurtosis is negative for both the variables which implies that both distributions have more rounded peak and shorter, thinner tails. The negative value of kurtosis also suggests that there is virtually no district in the country with extremely high index of age composition and districts with extremely high index of fertility.

The distribution of the sex ratio across the districts of the country has however been found to be negatively skewed in all the three indicators of female-male balance. This means that there are some districts in the country with extremely low proportion of females to males in the country. There are 9 districts in the country where the population sex ratio is estimated to be less than 800 females for every 1000 males according to the provisional figures of the 2011 population census with the lowest sex ratio estimated in district Daman in the Union Territory of Daman and Diu where there are only 533 females for every 1000 males. In district Leh of Jammu and Kashmir, also, the sex ratio has been estimated to be very low, just 583 females for every 1000 males. On the other hand, in 101 districts of the country, females outnumbered males at the 2011 population census with district Thane in Maharashtra topping the list with a population sex ratio of 1176 females for every 1000 males.

The negative skewness in the distribution across districts is sharper in case of the sex ratio in the population aged 0 to 6 years as compared to the population sex ratio. However, the lowest sex ratio in the population aged 0 to 6 years is well above the lowest population sex ratio. There are only six districts in the country where the sex ratio in the population aged 0 to 6 years has been estimated to be less than 800 females per 1000 males. Four of these six districts are in Haryana, while the remaining two are in Jammu and Kashmir. At the same time, there are only three districts in the country where the female population aged 0 to 6 years Outnumber the male population aged 0 to 6 years.

Finally, the distribution of the sex ratio in population aged 7 years and above across the districts is very much similar to the distribution of the population sex ratio. In fact, the sex ratio in the population aged 7 years and above appears to largely determine the population sex ratio. In ten districts of the country, the sex ratio in the population aged 7 years and above has been estimated to be less than 800 females per 1000 males and nine out of these ten districts, the population sex ratio is also less than 800 females per 1000 males with the lowest ratio estimated in district Daman of the Union Territory of Daman and Diu where there are only 500 females aged 7 years and above for every 1000 males aged 7 years and above. On the other hand, in 117 districts of the country, females aged 7 years and above outnumbered males 7 years and above with district Mahe in Puducherry, topping the list with more than 1200 females aged 7 years and above for every 1000 males 7 years and above.

Table 1 also presents estimates of the diversity index defined by Equation (4) for the country separately for the six demographic variables and summary measures of the inter-district variations in the district diversity index defined by Equation (3). The kernel density plots of the distribution of the diversity index across the districts are as shown in Figure 2. The district level diversity in the country is the highest in case of population density but lowest in case of the sex ratio in population aged 0 to 6 years, while the diversity in the population sex ratio lies between the diversity in sex ratio of the population aged 0 to 6 years and the diversity in the sex ratio in the population aged 7 years and above. On the other hand, the diversity in the index of fertility is higher than the diversity in the index of age composition. This implies that the distribution of the diversity index of the sex demographic variables across the districts of the country is essentially different. This observation is supported by the kernel density plots of the six demographic variables as shown in Figure 2. Notice that the kernel density plots of all the six variables are skewed towards the right because of the index of diversity defined by the Equation (3) is always positive irrespective of whether the intensity in the district is less than or higher than the national average.

In any case, it is evident from Table 1 and Figure 2 that the observed diversity across the districts in India in all the six demographic variables is primarily the result of extreme levels of diversity in selected districts. In most of the districts of the country, the diversity index, $D_{dc(v)}$, is less than the average values of $D_{dc(v)}$ for the country as a whole because of extreme diversity in selected districts.

The analysis suggests that when districts are put in the increasing order of the diversity index, 90% of the districts of the country account for only around 25% of the total diversity in population density across the districts of the country while the remaining 75% of the total diversity is accounted by the remaining 10% of the districts. In case of the index of age composition, this proportion is around 52% which means that nearly half of the diversity in the index of age composition across the districts of the country is accounted for by only 10% of the districts having extreme values of the index of age composition. In case of the population sex ratio, the corresponding proportions are 90:10:45:55. Similarly, the corresponding proportions are 90:10:51:49 in case of the sex ratio of the population aged 0 to 6 years and 90:10:44:66 in case of the sex ratio of the population aged 7 years and above. Finally, in case of the index of fertility, 90% of the districts of the country account for only about 51% of the total diversity across the districts while the remaining around

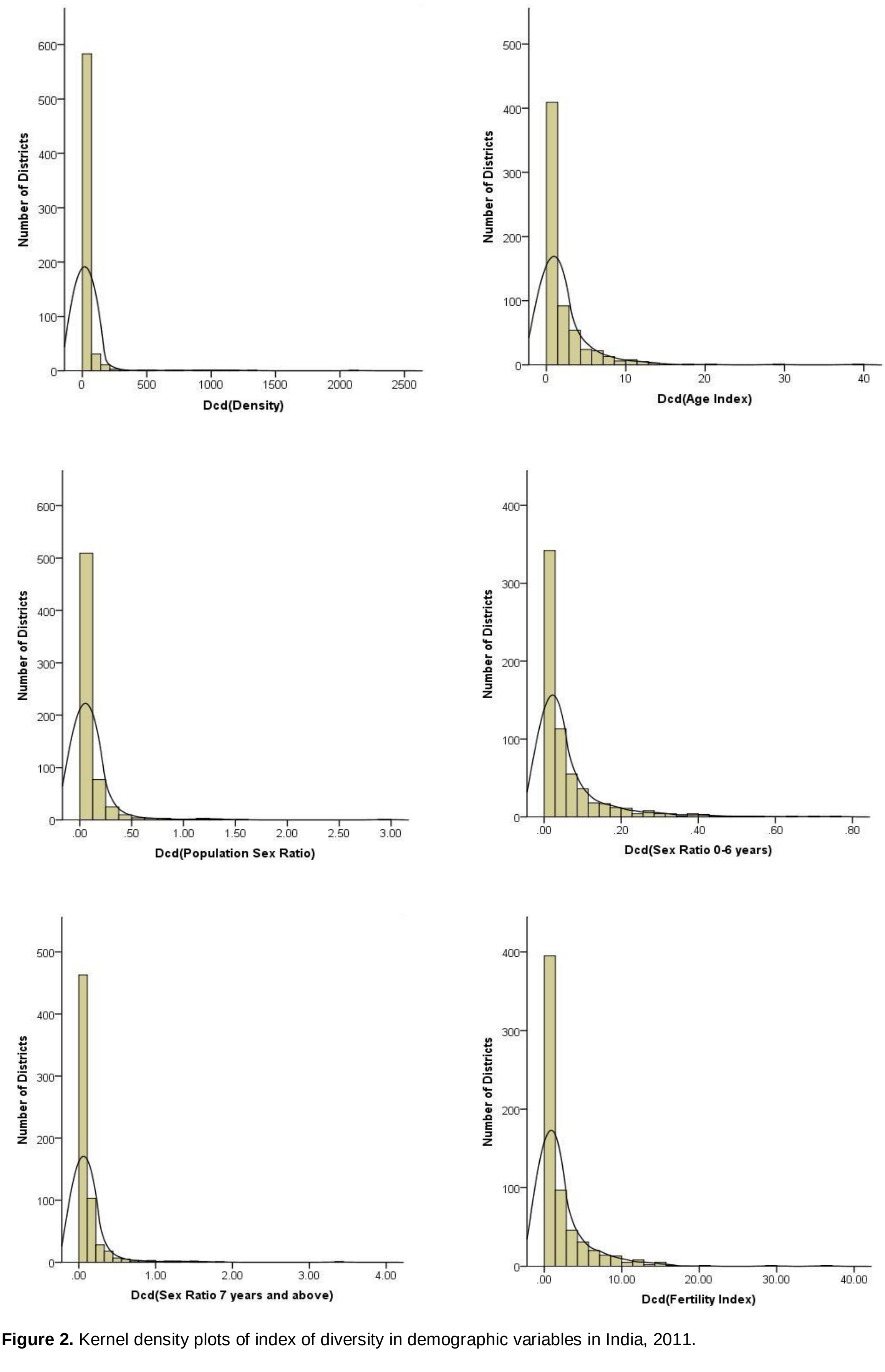

Figure 2. Kernel density plots of index of diversity in demographic variables in India, 2011.

Table 2. Ten districts with highest diversity in demographic variables.

Population density	Index of age composition	Population sex ratio	Sex ratio of the population aged 0 to 6 years	Sex ratio of the population aged 7 years and above	Index of fertility
1. Mumbai (Suburban)	1. Kolkata	1. Surat	1. Surat	1. Surat	1. Kolkata
2. Chennai	2. North 24-Parganas	2. Malappuram	2. Bid	2. Malappuram	2. North 24-Parganas
3. Kolkata	3. Mumbai (Suburban)	3. Kannur	3. Jalgaon	3. Kannur	3. Purba Champaran
4. Mumbai	4. Purba Champaran	4. Thrissur	4. Agra	4. Mumbai (Suburban)	4. Karimnagar
5. Delhi North-West	5. Karimnagar	5. Mumbai (Suburban)	5. Ahmednagar	5. Thrissur	5. Mumbai (Suburban)
6. Hyderabad	6. Hugli	6. Kollam	6. Sonipat	6. Kozhikode	6. Chennai
7. Bangalore	7. Chennai	7. Kozhikode	7. Jammu	7. Kollam	7. Hugli
8. Delhi North-East	8. Krishna	8. Tiruvananthapuram	8. Ahmedabad	8. Tiruvananthapuram	8. Krishna
9. Delhi West	9. Coimbatore	9. Daman	9. Amritsar	9. Daman	9. Coimbatore
10. North 24-Parganas	10. Mumbai	10. Alappuzha	10. Jhajhjhar	10. Thane	10. Tiruvananthapuram
Proportion (%) of the total diversity in the country explained by ten districts					
45.26	14.74	22.15	14.33	22.61	13.73

Author's calculations.

49% of the total diversity is accounted for by the remaining 10% of the districts. The highly skewed distribution of the districts of the country on the scale of the index of diversity in all the six demographic variables is very well reflected in the kernel density plots as shown in Figure 2. Table 2 lists 10 districts of the country with the highest diversity index $D_{dc(v)}$ for each of the six demographic variables. Table 2 also presents the proportion of the total diversity across the districts of the country accounted for by these ten districts. Results presented in Table 2 are revealing. In case of population density, 10 districts having the highest diversity index account for more than 45% of the total diversity across the districts of the country whereas in case of population sex ratio and the sex ratio of the population aged 7 years and above, this proportion is more than 22%. Finally, in case of the the index of age composition, sex ratio of the population aged 0 to 6 years and in case of the index of fertility, this proportion is around 14%. Table 2 confirms that a very substantial proportion of the total diversity in the selected demographic variables across the districts of the country as revealed through provisional figures of the 2011 population census is accounted for by extreme diversity in selected districts of the country. Table 2 also suggests that the districts with extreme diversity are different for different demographic variables, although there are districts which have extreme values in more than one variables included in the analysis. More specifically, district Mumbai Suburban in Maharashtra has extremely high diversity index in five of the six variables. It is only in case of the sex ratio of the population aged 0 to 6 years that this district is not included in the ten districtswith the highest diversity. On the other hand, district Chennai in Tamil Nadu and districts Kolkata and North 24 Parganas in West Bengal have extremely high diversity index in three of the six demographic variables: population density, index of age composition and the index of fertility. Similarly, district Tiruvananthapuram in Kerala has extremely high diversity index in three variables: population sex ratio, sex ratio of the population aged 7 years and above and the index of fertility, whereas district Surat in Gujarat has extremely high diversity index in all the three indicators of female-male balance in the population. In district Daman of the Union Territory of Daman and Diu and districts Kannur, Kollam, Kozhikode, Mallapuram and Thrissur of Kerala, the index of diversity has been found to be very high in the population sex ratio and in the sex ratio of the population aged 7 years and above. Similarly, in district Coimbatore of Tamil Nadu, districts Krishna and Karimnagar of Andhra Pradesh, district Hugli of West Bengal and district Purba Champaran of Bihar, extreme high diversity

Table 3. Cluster centres of the six demographic variables.

Variable	Cluster					
	1	2	3	4	5	6
Population density	10.860	127.037	2128.224	1216.664	508.966	916.094
Index of age composition	1.620	4.040	20.548	21.665	9.164	3.930
Population sex ratio	0.076	0.303	1.246	0.318	0.141	0.184
Sex ratio 0 to 6 years	0.054	0.106	0.003	0.108	0.147	0.107
Sex ratio 7+ years	0.086	0.358	1.544	0.359	0.159	0.221
Index of fertility	1.823	4.344	15.039	20.226	8.963	3.411
Number of districts in each cluster	583	45	1 Mumbai Suburban	3 Kolkata Mumbai Chennai	4 North 24 Parganas Haora Delhi South Delhi West	4 Hyderabad Bangalore Delhi North-East Delhi North-West

indexhas been estimated in case of the index of age composition and the index of fertility. In the remaining 16 districts, the diversity index is amongst the highest in the country in one of the six demographic variables. These districts include three districts of the National Capital Territory of Delhi: North-West, North-East and West districts, district Bangalore of Karnataka and district Hyderabad of Andhra Pradesh where the index of diversity is amongst the highest in the country in case of population density; districts Ahmednagar, Bid and Jalgaon of Maharashtra, districts Jhajhjhar and Sonipat of Haryana and district Amritsar in Punjab, district Jammu in Jammu and Kashmir and district Agra in Uttar Pradesh. Finally, district Thane in Maharashtra, the diversity index in the sex ratio of the population aged 7 years and above has been estimated to be amongst the highest across the districts of the country.

Thus, a very substantial proportion of the total diversity in the six demographic variables is accounted by only 34 or by just about 0.5% districts of the country. If these 34 districts are excluded from the analysis, then there is very substantial reduction in the inter-district diversity in the country of all the six demographic variables included in the analysis.

In order to cluster the districts according to the diversity in the six demographic variables, we have applied the K-means clustering technique. This technique requires the number of clusters to be extracted in advance. In order to decide the number of clusters, we used the rule of the thumb according to which the number of clusters should approximately be equal to $\sqrt{n/2}$, where n is the number of districts. This rule suggested that the total number of clusters to be extracted should be around 18.

The application of the aforementioned approach suggested that out of the 18 clusters so extracted, only five clusters have more than 2 districts in the cluster. In the remaining 13 clusters, there was only one district in 12 clusters and 2 districts in one cluster. The five clusters which have more than 2 districts, cluster one comprised of 464 or more than 72%; cluster two comprised of 103 or about 16% districts; cluster three comprised of 33 or about 5% districts, cluster four comprised of 18 or less than 3% districts and cluster five comprised of 8 or around 1% of the districts of the country. In other words, 567 or very close to 90% of the districts of the country are grouped in only two clusters in terms of the diversity in demographic situation depicted through the six demographic variables included in the analysis on the basis of the K-means clustering approach.

In view of the fact that using the rule of the thumb to decide about the number of clusters resulted in a large number of single district clusters, we reduced the number of clusters to six only and then applied the K-means clustering technique again to group the districts on the six dimensional scale of demographic situation. This exercise resulted in a very large cluster of 583 or more than 91% of the districts; a second cluster of 45 or around 7% of districts and four small clusters, two of which consist of four districts (Hyderabad, Bangalore, Delhi-North East, Delhi-

North West and North 24 Parganas, Haora, Delhi-South, Delhi-West), one consist of three districts (Kolkata, Mumbai and Chennai) and one only one districts (Mumbai Suburban). Results of the clustering exercise thus confirm that most of the demographic diversity across the districts of the country is largely due to some extreme diversity in less than 10% of the districts. In more than 90% of the districts of the country, the demographic diversity as reflected in terms of the six demographic variables used in the present analysis was quite small.

DECOMPOSITION

Districts of India are organised into States and Union Territories. This means that the diversity in the demographic variables across the districts of the country can be decomposed into inter-district diversity within the State/Union Territory and inter-State/Union Territory diversity according to Equation 13 for each district and according to Equation 14 for the whole country. Application of Equation 14 for individual States and Union Territories also permit the exploration of how the demographic diversity in different States/Union Territories contributes to the demographic diversity in the country as a whole.

Results of the exercise are presented in Table 4. Interestingly, the contribution of the within State/Union Territory and between State/Union Territory components of the inter-district diversity in the country varies by demographic variables used in the analysis. In case of population density, the within State/Union Territory component accounts for around 52% of the total inter-district diversity in the country.

Around 46% is accounted for between State/Union Territory component and the rest is accounted for, by the interaction term which is insignificant in terms of the magnitude. By comparison, the within State/Union Territory component accounts for around 28% of the inter-district diversity in the index of fertility while more than 71% of the diversity accounted between State and Union Territory component. It is also evident from the table that, except in the case of population density, the between State/Union Territory component is larger than the within State/Union Territory component in the remaining five variables.

As regards the relative contribution of different States/Union Territories to the total inter-district demographic diversity in the country, Uttar Pradesh figures amongst the five States contributing the largest share to the total diversity in all the six variables was included in the analysis. In addition to Uttar Pradesh, Andhra Pradesh, Maharashtra and Tamil Nadu figure in four of the six variables, Bihar and West Bengal figure in three variables, Kerala in two variables and Delhi, Haryana, Punjab and Rajasthan in one of the six demographic variables.

The total diversity accounted by the five States contributing the largest share has been estimated to be more than 55% in all but one demographic variable. It is only in case of the sex ratio of the population aged 0 to 6 years that the five States with largest contribution accounts for very close to 50% of the inter-district diversity in the country while this proportion is estimated to be more than 60% in case of the index of fertility. Table 4 also suggests that the relative contribution of within State/Union Territory and between State/Union Territory diversity varies by State/Union Territory for different demographic variables.

CONCLUSIONS

The analysis based on the provisional results of the 2011 population census suggests that most of the demographic diversity across the districts of the country as reflected in the six demographic variables included in the analysis is confined to a few districts only. This observation bears significance in view of the social, cultural, economic and ecological diversity that is so pervasive in India. Despite this social, cultural, economic and ecological diversity, the demographic scenario appears very similar across most of the districts of the country and there are only a few districts with extreme diversity in the demographic situation.

The demographic diversity in the six demographic variables used in the present analysis may be attributed to the interaction of basic demographic processes: fertility, mortality and migration. Information about these demographic processes at the district level is currently not available through the 2011 population census. Once detailed district level information about fertility, mortality and migration is available from the 2011 population census, it will be possible to analyse in greater detail the factors responsible for demographic similarity or demographic diversity across the districts of the country.

The analysis also suggests that between State diversity in the six demographic variables accounts for a larger proportion of the total inter-district diversity in the country as compared to within State inter-district diversity and there are substantive inter-State variations. Demographic diversity across the States/Union Territories of India is well known. The provisional results of the2011 population census suggest that this diversity continue to persist.

Finally, a caveat, the analysis presented here is based on the provisional figures of the 2011 population census. These figures are likely to be changed when final figures of the 2011 population census are released. The relevance of the present analysis lies in the fact that it will take almost a year when final figures of the population census are released. There is however little probability of any significant difference between the provisional figures used in this analysis and final figures to be released by

Table 4. Within State/Union Territory and between States/Union Territories components of the diversity in demographic variables across the districts in India.

State/Union Territory	Population density				Index of age composition			
	Within State/UT	Between States/UTs	Interaction	Total	Within State/UT	Between States/UTs	Interaction	Total
AN Islands	1.680	25.910	-3.550	24.036	0.023	0.412	0.005	0.440
Andhra Pradesh	1200.300	60.370	-116.380	1144.290	13.954	106.441	3.163	123.558
Arunachal Pradesh	10.840	212.430	-39.110	184.163	0.654	0.360	-0.059	0.954
Assam	183.320	0.890	8.360	192.565	18.767	6.291	-1.492	23.566
Bihar	209.530	1656.140	259.300	2124.964	12.989	218.194	-3.061	228.122
Chandigarh	0.000	167.180	0.000	167.179	0.000	0.543	0.000	0.543
Chhattisgarh	116.380	196.340	-94.940	217.769	3.625	2.406	-0.203	5.828
Dadra and Nagar Haveli	0.000	1.960	0.000	1.957	0.000	0.057	0.000	0.057
Daman and Diu	0.330	11.440	0.630	12.408	0.024	0.211	0.005	0.240
Delhi	181.810	2996.220	553.770	3731.794	1.402	4.071	0.133	5.606
Goa	0.770	0.020	0.020	0.819	0.057	2.875	0.016	2.948
Gujarat	607.480	42.340	-130.530	519.282	25.235	3.828	1.243	30.305
Haryana	38.920	42.800	12.900	94.624	10.854	0.040	0.085	10.978
Himachal Pradesh	72.150	136.440	-123.580	85.019	0.939	3.703	0.137	4.779
Jammu and Kashmir	374.990	244.400	-466.550	152.836	15.646	10.578	-2.407	23.817
Jharkhand	137.230	3.420	10.550	151.205	12.659	25.710	-1.917	36.452
Karnataka	1109.020	30.470	-107.850	1031.648	31.501	30.412	4.386	66.299
Kerala	94.320	347.170	88.790	530.275	14.332	50.793	3.662	68.788
Lakshadweep	0.000	2.780	0.000	2.781	0.000	0.040	0.000	0.040
Madhya Pradesh	260.220	262.770	-103.710	419.269	21.701	15.845	-1.826	35.719
Maharashtra	3753.920	3.260	-70.490	3686.683	37.962	43.105	4.542	85.609
Manipur	97.560	55.190	-77.700	75.047	0.643	0.007	0.006	0.656
Meghalaya	12.580	51.570	-11.900	52.239	1.052	8.310	-0.279	9.083
Mizoram	4.990	67.630	-8.690	63.925	0.430	0.487	-0.051	0.866
Nagaland	14.680	40.990	-14.500	41.170	1.197	0.380	-0.093	1.484
Orissa	244.680	79.060	-76.850	246.886	20.377	6.759	1.584	28.720
Puducherry	2.760	69.450	5.850	78.053	0.021	1.510	0.005	1.536
Punjab	54.440	58.030	18.600	131.060	1.348	24.889	0.257	26.494
Rajasthan	472.570	444.300	-327.760	589.106	16.480	34.537	-2.057	48.961
Sikkim	11.510	21.120	-19.520	13.107	0.043	0.860	0.010	0.913
Tamil Nadu	1237.960	159.890	212.310	1610.161	8.243	143.341	2.369	153.953
Tripura	12.550	0.420	-1.140	11.836	1.028	0.495	0.073	1.596
Uttar Pradesh	563.930	1874.460	446.320	2884.710	45.222	67.642	-4.617	108.247
Uttarakhand	161.060	77.790	-99.040	139.807	1.360	0.000	-0.001	1.359
West Bengal	1081.320	1388.740	643.660	3113.716	63.282	53.175	9.311	125.768

Table 4. Contd.

State/Union Territory	Population sex ratio				Sex ratio of the population aged 0 to 6 years			
	Within State/UT	Between States/UTs	Interaction	Total	Within State/UT	Between States/UTs	Interaction	Total
India	12325.800	10833.390	367.270	23526.460	383.050	868.307	12.929	1264.286
	52.390	46.050	1.560	100.000	30.300	68.680	1.020	100.000
AN Islands	0.013	0.028	0.000	0.041	0.000	0.018	-0.000	0.018
Andhra Pradesh	0.543	3.737	-0.000	4.280	0.346	1.259	0.032	1.637
Arunachal Pradesh	0.164	0.011	0.000	0.175	0.009	0.052	0.001	0.062
Assam	0.069	0.096	-0.000	0.165	0.076	1.021	0.053	1.150
Bihar	1.251	1.131	0.001	2.383	0.662	0.639	-0.036	1.265
Chandigarh	0.000	0.322	0.000	0.322	0.000	0.047	0.000	0.047
Chhattisgarh	0.114	1.099	-0.000	1.212	0.069	1.113	-0.011	1.171
Dadra and Nagar Haveli	0.000	0.200	0.000	0.200	0.000	0.001	0.000	0.001
Daman and Diu	0.276	0.665	0.021	0.962	0.000	0.000	0.000	0.000
Delhi	0.099	1.750	0.001	1.849	0.072	0.751	0.001	0.824
Goa	0.003	0.019	-0.000	0.022	0.003	0.001	-0.000	0.003
Gujarat	2.896	0.540	0.004	3.440	1.069	0.948	0.146	2.162
Haryana	0.115	1.899	0.001	2.015	0.462	3.692	0.284	4.438
Himachal Pradesh	0.372	0.131	-0.000	0.502	0.116	0.008	0.002	0.126
Jammu and Kashmir	0.634	0.782	0.004	1.420	0.442	0.767	0.060	1.270
Jharkhand	0.380	0.026	-0.000	0.406	0.197	0.486	-0.052	0.631
Karnataka	1.183	0.822	-0.001	2.004	0.116	0.936	0.048	1.100
Kerala	0.517	10.523	0.003	11.043	0.020	1.187	-0.011	1.195
Lakshadweep	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Madhya Pradesh	2.257	0.133	0.001	2.391	1.468	0.007	0.002	1.477
Maharashtra	3.775	0.441	0.002	4.218	2.786	2.074	-0.269	4.591
Manipur	0.051	0.100	-0.000	0.151	0.008	0.019	-0.001	0.026
Meghalaya	0.021	0.103	-0.000	0.124	0.003	0.160	-0.001	0.162
Mizoram	0.016	0.023	-0.000	0.039	0.005	0.063	0.001	0.069
Nagaland	0.020	0.003	0.000	0.023	0.027	0.032	0.006	0.065
Orissa	0.619	1.038	-0.000	1.657	0.643	0.311	-0.101	0.854
Puducherry	0.011	0.191	0.000	0.203	0.003	0.057	0.000	0.060
Punjab	0.292	1.147	0.001	1.440	0.196	2.577	-0.039	2.735
Rajasthan	1.536	0.234	0.001	1.771	0.838	1.332	0.085	2.255
Sikkim	0.025	0.030	0.000	0.054	0.002	0.010	0.000	0.012
Tamil Nadu	0.470	3.606	-0.000	4.076	0.513	1.345	0.037	1.894
Tripura	0.003	0.028	-0.000	0.030	0.010	0.099	-0.007	0.102

Table 4. Contd.

Uttar Pradesh	7.216	3.767	0.010	10.993	3.140	0.857	0.083	4.080
Uttarakhand	1.233	0.089	-0.000	1.321	0.098	0.161	-0.006	0.253
West Bengal	0.302	0.064	-0.000	0.365	0.095	2.041	-0.062	2.074
India	26.476	34.778	0.049	61.303	13.494	24.071	0.245	37.810
	43.190	56.730	0.080	100.000	35.690	63.660	0.650	100.000

State/Union Territory	Sex ratio of the population aged 7 years and above				Index of fertility			
	Within State/UT	Between States/UTs	Interaction	Total	Within State/UT	Between States/UTs	Interaction	Total
AN Islands	0.016	0.042	0.000	0.058	0.030	0.286	0.003	0.319
Andhra Pradesh	0.627	3.950	-0.016	4.561	15.515	128.270	3.378	147.162
Arunachal Pradesh	0.218	0.025	-0.001	0.242	0.587	0.464	-0.061	0.990
Assam	0.098	0.041	0.001	0.141	18.297	5.780	-1.434	22.644
Bihar	1.646	1.943	0.014	3.603	14.370	240.076	-2.973	251.473
Chandigarh	0.000	0.377	0.000	0.377	0.000	0.168	0.000	0.168
Chhattisgarh	0.133	1.111	-0.003	1.241	3.855	1.032	-0.130	4.757
Dadra and Nagar Haveli	0.000	0.277	0.000	0.277	0.000	0.276	0.000	0.276
Daman and Diu	0.347	0.841	0.022	1.210	0.035	0.004	0.003	0.043
Delhi	0.109	1.937	-0.001	2.045	1.337	1.656	0.083	3.077
Goa	0.003	0.021	0.000	0.024	0.044	3.130	0.017	3.191
Gujarat	3.389	0.503	-0.010	3.881	23.922	2.527	0.978	27.426
Haryana	0.135	1.685	-0.003	1.816	10.731	0.229	-0.203	10.757
Himachal Pradesh	0.492	0.168	-0.001	0.660	1.139	4.551	0.124	5.814
Jammu and Kashmir	0.855	0.758	-0.011	1.601	14.684	13.740	-2.713	25.712
Jharkhand	0.456	0.007	0.001	0.464	12.248	25.269	-1.892	35.624
Karnataka	1.423	0.770	-0.003	2.190	32.400	35.555	4.664	72.619
Kerala	0.646	11.948	0.028	12.622	13.755	78.134	4.514	96.403
Lakshadweep	0.000	0.000	0.000	0.000	0.000	0.042	0.000	0.042
Madhya Pradesh	2.538	0.156	0.000	2.694	22.342	17.508	-1.839	38.011
Maharashtra	4.630	0.348	-0.005	4.973	36.668	39.201	4.184	80.052
Manipur	0.063	0.117	0.000	0.180	0.643	0.065	0.018	0.727
Meghalaya	0.033	0.101	-0.001	0.133	1.094	7.404	-0.258	8.241
Mizoram	0.022	0.019	-0.001	0.039	0.510	0.395	-0.043	0.861
Nagaland	0.025	0.008	-0.000	0.033	1.198	0.441	-0.100	1.540
Orissa	0.711	1.147	0.030	1.888	17.682	9.892	1.938	29.512
Puducherry	0.015	0.208	0.000	0.223	0.013	2.123	0.005	2.142
Punjab	0.347	1.054	0.002	1.403	1.592	19.843	0.206	21.641
Rajasthan	1.894	0.112	0.004	2.011	17.959	36.599	-1.994	52.565

Table 4. Contd.

Sikkim	0.029	0.042	-0.000	0.071	0.033	0.672	0.008	0.713
Tamil Nadu	0.518	3.742	-0.003	4.256	8.669	167.792	2.502	178.963
Tripura	0.004	0.020	-0.000	0.024	1.070	0.604	0.080	1.753
Uttar Pradesh	8.953	4.330	0.036	13.320	49.439	86.574	-4.766	131.247
Uttarakhand	1.586	0.163	-0.004	1.745	2.169	0.036	0.008	2.213
West Bengal	0.344	0.006	0.001	0.350	61.490	53.759	9.376	124.625
India	32.305	37.977	0.076	70.358	385.520	984.097	13.683	1383.300
	45.920	53.970	0.110	100.000	27.870	71.140	0.990	100.000

UTs: Union Territories.

the Registrar General and Census Commissioner of India.

REFERENCES

Durand JD (1948). The Labour Force in the United States: 1890-1960. Appendix B: Methods of Analyzing Labour Force Change. New York, Social Sci. Res. Council, p. 288.

Government of India (2000). National Population Policy 2000. New Delhi, Ministry of Health and Family Welfare, p. 88

Government of India (2002). Xth Five-year Development Plan 2002-2007. I: Dimensions and Strategies. New Delhi, Planning Commission, p. 232.

Government of India (2002). National Health Policy 2002. New Delhi, Ministry of Health and Family Welfare. p. 108.

Government of India (2005). National Rural Health Mission. New Delhi, Ministry of Health and Family Welfare. p. 80.

Government of India (2008) Eleventh Five-year Development Plan 2007-2012. Volume I: Inclusive Growth. New Delhi, Planning Commission, p. 306.

Government of India (2011). Census of India 2011. Provisional Population Totals. Paper 1 of 2011. India, Series 1. New Delhi, Registrar General and Census Commissioner, p.188.