STUDIES

No. 13-07

Workers in Villages of Gujarat

Aalok Ranjan Chaurasia

'Shyam' Institute

This page is intentionally left blank

WORKERS IN VILLAGES OF GUJARAT

INTRODUCTION

The only source of data about the size and structure of work force at the village level is the population census. On the basis of the information available through the 2011 population census, work participation rate (WPR) has been calculated for all villages of the State for the total population and separately for males and females. Work participation rate has also been calculated out for different categories of workers. The work classification used at the 2011 population census allows to calculate 33 indicators of work participation for every village of the State. This exercise provides vary valuable information about the village economy in the State as participation rate in the productive activity is a reflection of the state of the economy.

In the Indian population census, classification of a habitation as a village is essentially a residual concept. At the 2011 population census, all habitations were classified into urban and rural habitations on the basis of the definition of an urban habitation. All habitations which do not conform to the definition of an urban habitation are classified as rural habitation. The basic unit of the rural habitation in the 2011 population census was the revenue village as recognised by the local administration. The revenue village need not necessarily be a single agglomeration of the habitations. The revenue village however has a definite surveyed boundary and each village is a separate administrative unit with separate village accounts. A village may have one or more hamlets. The entire revenue village is one unit. There may be un-surveyed villages such as villages within the forests where locally

recognised boundaries of each habitation area is followed within the larger unit of say the forest range. Population living within the boundaries of the revenue village is taken as the population of the village. There may be a possibility that there is no population in a revenue village. Similarly, a revenue village may be inhabited by only males or by only females or by just one person or a few persons.

According to the 2011 population census, there were 17843 village in Gujarat with varying population size. Five villages had only one habitant at the time of census whereas in 249 villages, the total number inhabitants was less than 100. Similarly, there were 11 villages where there was no female population. On the other hand, there were 1127 villages which had a population of more than 5000 inhabitants with the village Dhari in sub-district Dhari of district Amreli was the only village in the State which had a population of more than 30 thousand at the 2011 population census. The average population size of a village in the State was 1945 inhabitants while the median population size was 1390. The large difference between the mean and median population size of a village in the State suggests that the distribution of the villages of the State by the size of the population is highly positively skewed. The coefficient of skewness of the distribution of villages by the enumerated population is 3.06 which implies that the number of villages having population less than the State average is substantially higher than the number of villages having population above the State average. In other words, majority of the villages in the State are small villages with population less than 2000 inhabitants at the time of the 2011 population census. At the same time, there are some very large villages in the State which had not been classified as urban areas at the 2011 population census as the population composition of these village do not conform to the definition of an urban area adopted at the 2011 population census.

METHOD OF ANALYSIS AND RESULTS

The analytical approach adopted here consists of characterising villages on the basis of the work participation rate. The underlying assumption of this characterisation is that the work participation rate across the villages is an independent and identically distributed random variable. This essentially means that level of work participation rate in one village is not influenced by the work participation rate in other villages. Moreover, it is also assumed that the distribution of the work participation rate across villages does not follow a pre-determined pattern.

The characterisation of the distribution of villages on the basis of the work participation rate may be done by following either a parametric approach or a non-parametric approach. The parametric approach generally assumes that the underlying unknown distribution or variation of the work participation rate across villages is essentially normal so that estimates of arithmetic mean and standard deviation of the work participation rate are sufficient to characterise the distribution

of villages by the work participation rate. This approach does not yield desired results if the underlying distribution of villages by the work force participation rate is not normal.

The non-parametric approach, on the other hand, makes no assumption about the underlying unknown distribution. In this approach, the contribution of each village is smoothed out from a single point into a region or space surrounding it. Aggregating the individually smoothed contributions of all villages gives an overall picture of the distribution of the work participation rate across villages and associated density function.

Let $(x_1, x_2, \ldots, x_n)$ is a sample drawn from some distribution with an unknown density f. The density function f can be estimated by the kernel density estimator

$$\hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i) \quad = \frac{1}{nh}\sum_{i=1}^{n} K\Big(\frac{x - x_i}{h}\Big),$$

where K is a symmetric but not necessarily positive function that integrates to one and has mean zero and $h > 0$ is a smoothing parameter called the bandwidth. A kernel with subscript h is called the scaled kernel and defined as $K_h(x) = 1/h\, K(x/h)$. The smoothness of the density curve depends upon the value of h. The higher is the value of h, the smoother is the density function. Intuitively h should be chosen as small as the data allow. However there is always a trade-off between the bias of the estimator and its variance. On the other hand, selection of the kernel function is not very important. In the present analysis, we have used the Epanechnikov kernel which is optimal in a minimum variance sense.

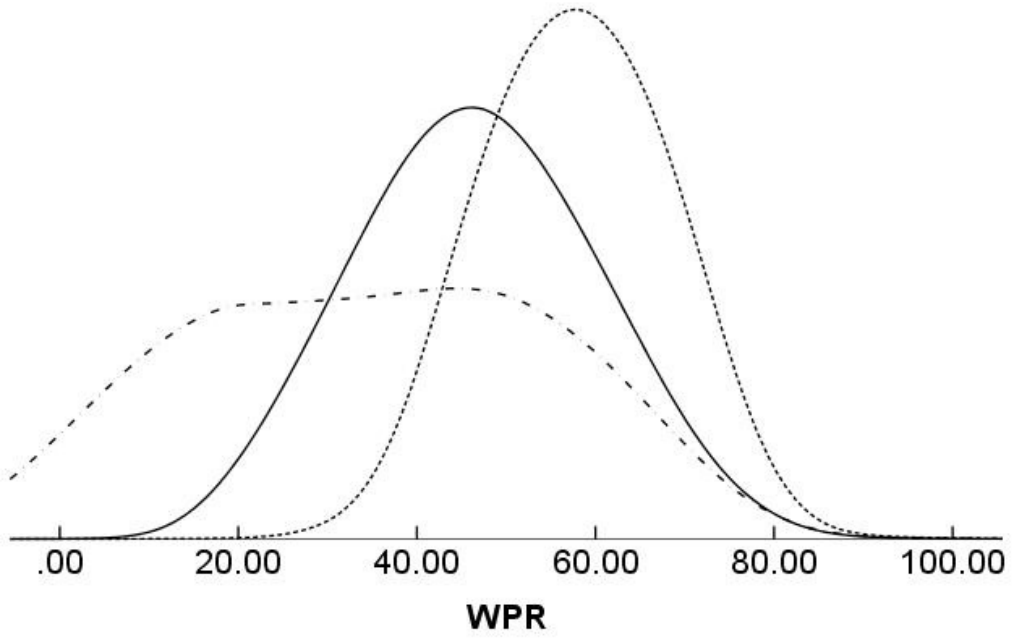
The work participation rates used in the present analysis have been defined as follows:

Work participation rate	The proportion (per cent) of workers to the total population.
Main work participation rate	The proportion (per cent) of main workers to the total population.
Marginal work participation	The proportion (per cent) of marginal workers to the total population.

Main and marginal workers are further divided into cultivators, agricultural labourers, household industry workers and other workers. Accordingly, the participation rate has been calculated for 33 categories of workers for each village. Female work participation rate, however, could be calculated for 17832 villages as there was no female population in 11 villages.

Figure 1
Distribution of work participation rate in villages of Gujarat

_____ Total - - - Male - - - Female



The first step in the analysis of the distribution of the villages of the State by 36 work participation rates is to calculate the parameters of the distribution. Statistically, the parameters that characterise any distribution can be grouped into the following four categories:

1. Measures of central tendency, the most common of which are arithmetic mean and median.

2. Measures of dispersion, the most common of which are standard deviation, variance and the inter-quartile range.

3. Measure of deviation of the distribution from normality or the coefficient of skewness. In case of a normal probability density function, the coefficient of skewness is zero.

4. Measure of the concentration of the probability density function around central tendency which is measured in terms of kurtosis. For a normal distribution, the kurtosis is 3. A value of kurtosis higher than 3 makes the

distribution leptokurtic whereas a value of kurtosis less than 3 makes the distribution platokurtic. A leptokurtic distribution has long peak but flatter tails. A platokurtic distribution has wider peak around central tendency but thinner tails.

Table 1 presents the summary measures of the distribution of 33 work participation rates across the villages of the State. Interestingly, in a number of categories of workers, the work participation rate varies from a minimum of 0 per cent to a maximum of 100 per cent. A work participation rate of 100 per cent in a village means that all the inhabitants of the village are classified as workers at the time of the population census. This actually means that the entire population of the village is the working age population. This situation may be possible villages are actually revenue villages. In very small villages, extremely high or extremely low work participation rates are possible because the population census in India is carried out on the de facto basis so that people working in fields get enumerated in the revenue village in which the fields are located. Population census in India is not conducted on a de jure basis.

One way to exclude villages with extreme level of work participation rates is to estimate 5 per cent trimmed mean. Trimmed mean does not take into account extreme values. If the extreme values of work participation rates are excluded from the analysis then the mean work participation rates across the villages of the State is estimated to be 46.50 per cent – 57.81 per cent for males and 34.76 per cent for females. On the other hand, the median is estimated to be 46.51 - 57.85 per cent for males and 35.06 for females. The median is not influenced by the extreme values.

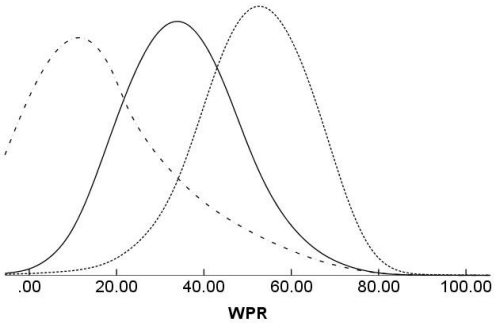
It may be seen from the table that arithmetic mean and median of inter-village distribution of the work participation rate are very close. This implies that the probability density function closely resembles the normal distribution. They would have been the same if the probability density function of the inter-village distribution of the work participation rateis exactly normal.

The very fact that the distribution of the work participation rate across the villages of the State is very close to a normal distribution is also reflected through the coefficient of skewness which is very close to zero for the total population and for males and females separately. A normal distribution of the work participation rates across villages implies that the proportion of villages having work participation rate lower than the arithmetic mean is the same as the proportion of villages having work participation rate higher than the arithmetic mean.

Figure 2
Participation rate of different categories of work

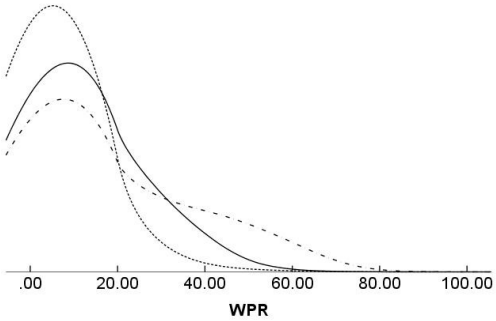
Main workers

All categories

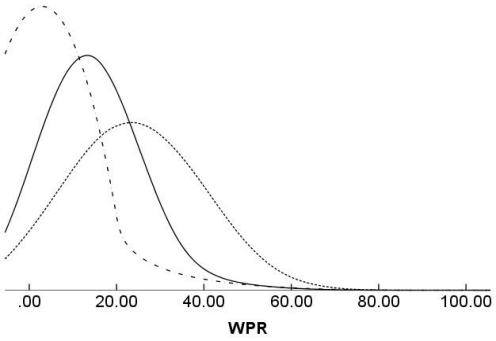


Marginal workers

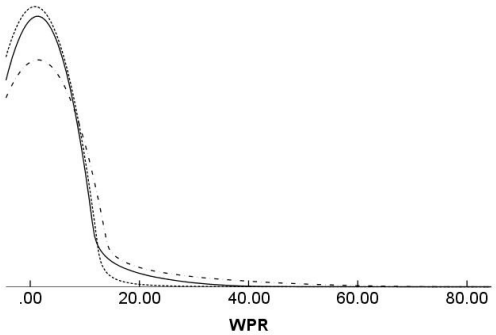
All categories



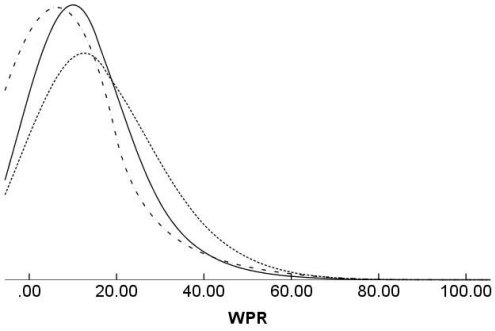
Cultivators



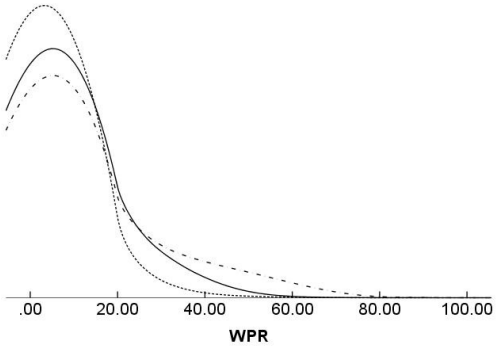
Cultivators



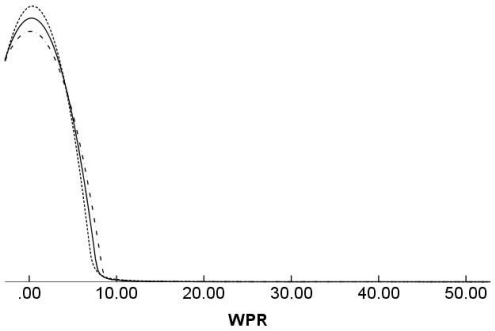
Agricultural labourers



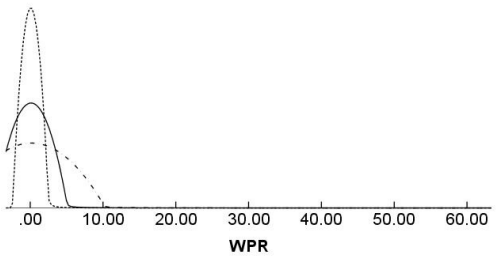
Agricultural Labourers



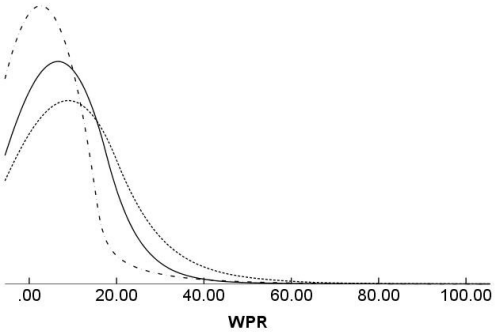
Household industry worker



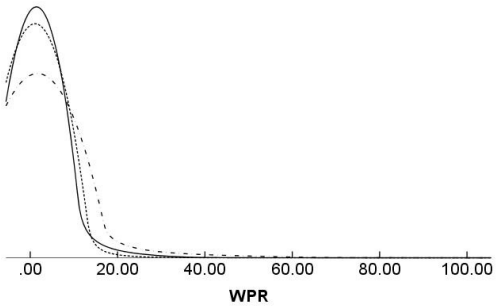
Household industry worker



Other workers



Other workers



On the other hand, the dispersion of villages from the central tendency in terms of the work participation rate is comparatively very low in males than in females whereas the coefficient of kurtosis is very highly positive for male work participation rate but very highly negative for the female work participation rate. This means that in most of the villages of the State, the male work participation rate is very close to the average male work participation rate of all villages but the same is not true for female work participation rates. In other words, the distribution of male work participation rate across districts is essentially leptokurtic with long thin peak and long tails whereas the distribution of female work participation rate across villages is in fact platokurtic with flat broad peak and short tails. In other words, the probability density function of the distribution of the male work participation rate across villages is essentially different from the distribution of female work participation rate across villages.

The above observations are brought out more clearly through the kernel density plot which has been presented in figure 1. The kernel density plots confirm that the distribution of male work participation rate is essentially different from the distribution of female work participation rate across the villages of the State. The male work participation rate varies in a close range across villages but the same is not the case with the female work participation rate. As a result, the inter-quartile range for the male work participation rate is only 7.92 but the same for the female work participation rate 32.75 which is almost four times higher than the male work participation rate. In fact, in almost two-third villages of the State, the male work participation rate varies between 40-60 per cent whereas the female work participation rate is less than 20 per cent in around 27 per cent of villages; ranges between 20-40 per cent in almost 30 per cent of villages and is between 40-60 per cent in about one third of the villages of the State (Table 2). It may also be seen from the figure 1 that the kernel density plot of the male and female combined work participation rate is a juxtaposition of the kernel density plot of male and female work participation rate and reflects a near normal probability density function.

Compared to the work participation rate, the distribution of main work participation rate across villages is different for the combined population as well as for males and females. The first observation that can be made from the kernel density plot (Figure 2) is that the probability density function of the distribution of main work participation rate across villages is skewed and the skewness in the distribution is relatively more in the female main work participation rate as compared to the male main work participation rate. This fact is also supported by the difference in the mean and median main work participation rate and the coefficient of skewness which is large in female main work participation rate than in male main work participation rate. For the combined population, the coefficient of skewness is positive which means that the proportion of villages with main work participation rate lower than the average is higher than the proportion of villages with main work participation rate higher than the average. The same interpretation holds for the inter-village distribution of female main work participation rate.

However the shape of the distribution is different in case of male main work participation rate. Here the proportion of villages having male main work participation rate higher than the average is more than the proportion of villages having male main work participation rate lower than the average. On the other hand, contrary to the main work participation rate, the coefficient of kurtosis of the distribution of main work participation rate across villages is nearly the same for the male main workers, female main workers and male and female combined main workers.

The kernel density plots of marginal work participation rate are presented in figure 3 for the combined population and separately for males and females. It can be seen from the figure that male marginal work participation rate is heavily positively skewed compared to the female marginal work participation rate – the coefficient of skewness of the distribution of male marginal work participation rate is 2.60 which is more than 2.5 times the coefficient of skewness of the distribution of female marginal work participation rate. The skewness in the distribution is also reflected in the different between the mean and the median marginal work participation rate. Moreover the marginal work participation rate is substantially lower than the main work participation rate in the combined population as well as separately in males and females and between males and females, it is substantially higher in females as compared to males. The kernel density plots presented in figure 3 suggest there are some villages in the State where the marginal work participation rate is substantially high.

The kernel density plots of the distribution of other work participation rates across the villages of the State calculated from the 2011 population census data are presented in figures 4 through 11 while summary measures of the distribution are presented in table 1. It is obvious from the figures and the table that inter-village distribution of different work participation rates is different and the participation rates are the lowest in case of household industry but the highest in farming.

It would be interesting to analyse whether the work participation rate varies by the population size of the village. This analysis bears importance in view of the fact that the larger is the village in terms of population size, the larger is its economy so that the work participation rate may vary by the size of the population of the village. To analyse this relationship, we have grouped the villages of the State into eight categories according to the size of the population of the village as enumerated at the 2011 population census as shown in table 3 and, for each group, calculated the 33 indicators related to the work participation. It may be seen from the table that the total work participation rate does no vary much across the eight categories of villages as the estimated coefficient of variation is quite small. This means that the overall work participation rate does not vary much by the size of the population of the village. A similar situation appears to prevail in case of male and female work participation rates as well as in case of main and marginal work participation rates where the coefficient of variation is always less than 15 per cent.

The household industry workers participation rate and other workers participation rate have however been found to vary by the size of the population of the village and these rates are the highest in villages with at least 10 thousand population at the 2011 population census. In other words, the size of the village economy has a direct impact on the participation rates in terms of household industry sector and in terms of other workers.

CONCLUSIONS

The present analysis is probably the first to analyse the work participation rate at the village level on the basis of the data collected at the 2011 population census. Since the work participation rate is a reflection of the state of economy, the analysis also provides some valuable insight about the village economy. It may be reiterated here that the data collected during the population census is the only source of information about the work status of the people at the village level and about the village economy.

Key findings of the analysis presented in this paper may be summarised as under:

1. The work participation rate in the villages of the State are dominated by the work in the primary sector of the economy – cultivation and agriculture labour. This means the village economy of the State depends, almost entirely, upon the primary sector of the economy, particularly the agriculture. Other sectors of the economy have only a small presence at the village level.

2. Even in very large villages – villages with a population of at least 10 thousand at the 2011 population census, sectors other than the primary sector account for only a small proportion of the work force.

3. The distribution of all sectors combined work participation rate across the villages of the State is distributed normally but this is not the case with sector specific work participation rates and there are important gender differentials. In most of the cases, the distribution of work participation rate across villages is positively skewed which implies that in most of the villages, the work participation rate is lower than the village average.

4. It would be interesting to carry out a statistical clustering exercise to group villages on the basis of the structure of the work force. Such a clustering exercise may help in characterising the village economy of the State. This exercise may provide further insight to the village economy.

Table 1
Summary measures of the distribution of work participation rates across villages of Gujarat, 2011

Participation rate	Mean	Trimmed mean	Median	Standard deviation	IQR	Skewness	Kurtosis
All workers	46.64	46.50	46.51	10.73	15.99	0.19	-0.26
Male workers	57.79	57.81	57.85	6.56	7.92	0.00	3.24
Female workers	34.86	34.76	35.06	19.27	32.75	0.01	-1.06
Main workers -All	35.00	34.70	33.73	10.65	12.86	0.49	0.76
Main workers - Male	51.64	52.14	53.09	9.82	11.22	-0.99	2.83
Main workers - Female	17.50	16.16	12.56	15.68	21.09	1.15	0.71
Marginal Workers - All	11.64	10.64	7.62	11.72	16.97	1.15	0.97
Marginal Workers - Male	6.15	5.13	3.30	7.79	7.26	2.60	10.37
Marginal Workers - Female	17.36	15.93	10.18	18.39	27.74	0.98	-0.15
Main workers – Cultivator - All	14.52	13.92	13.53	9.02	11.01	1.18	2.92
Main workers – Cultivator - Male	24.12	23.88	23.98	12.75	18.52	0.23	-0.31
Main workers – Cultivator - Female	4.49	3.05	1.12	8.53	3.53	3.35	13.36
Main workers – Agricultural labourers - All	12.78	11.85	10.69	10.55	12.82	1.37	2.56
Main workers – Agricultural labourers - Male	16.30	15.42	14.19	12.58	17.12	0.94	0.77
Main workers – Agricultural labourers - Female	9.06	7.59	4.76	11.29	11.01	2.08	4.94
Main workers – Household industry workers - All	0.29	0.17	0.07	0.83	0.29	13.66	376.09
Main workers – Household industry workers - Male	0.42	0.26	0.09	1.06	0.44	10.03	190.7
Main workers – Household industry workers - Female	0.16	0.05	0.00	0.93	0.08	22.09	715.56
Main workers – Others - All	7.41	6.63	5.49	7.17	7.77	2.39	10.79
Main workers – Others - Male	10.80	9.65	7.93	10.39	11.32	2.07	6.81
Main workers – Others - Female	3.79	2.64	1.49	6.84	2.93	4.11	21.71

Participation rate	Mean	Trimmed mean	Median	Standard deviation	IQR	Skewness	Kurtosis
Marginal workers – Cultivator - All	2.20	1.37	0.39	4.73	1.61	3.68	16.58
Marginal workers – Cultivator - Male	0.98	0.63	0.30	2.25	0.95	7.46	99.61
Marginal workers – Cultivator - Female	3.46	1.96	0.32	8.23	1.95	3.60	14.48
Marginal workers – Agricultural labourers - All	7.33	6.13	3.06	9.8	9.90	1.89	3.95
Marginal workers – Agricultural labourers - Male	3.87	2.91	1.29	6.35	4.40	3.35	17.63
Marginal workers – Agricultural labourers - Female	10.94	9.08	3.78	15.18	14.96	1.72	2.3
Marginal workers – Household industry workers - All	0.14	0.06	0.00	0.64	0.08	14.96	347.72
Marginal workers – Household industry workers - Male	0.09	0.03	0.00	0.35	0.00	12.72	260.17
Marginal workers – Household industry workers - Female	0.20	0.06	0.00	1.12	0.08	18.02	505.66
Marginal workers – Others – All	1.97	1.30	0.58	3.91	1.70	4.18	23.34
Marginal workers – Others - Male	1.21	0.84	0.47	2.39	1.23	6.16	71.65
Marginal workers – Others - Female	2.77	1.53	0.47	6.8	1.88	4.61	26.22

Table 2
Distribution of villages by work participation rate

WPR	Number	Percent	Number	Percent	Number	Percent
			All workers			
42388	26	0.1	8	0.0	4875	27.3
20-40	5328	29.9	115	0.6	5343	29.9
40-60	10426	58.4	11444	64.1	5825	32.6
60-80	2023	11.3	6210	34.8	1757	9.8
80-100	40	0.2	66	0.4	32	0.2
All	17843	100.0	17843	100.0	17832	99.9
0-20	1038	5.8	179	1.0	11718	65.7
20-40	11860	66.5	1909	10.7	4158	23.3
40-60	4553	25.5	12792	71.7	1631	9.1
60-80	374	2.1	2926	16.4	320	1.8
80-100	18	0.1	37	0.2	5	0.0
All	17843	100.0	17843	100.0	17832	99.9
0-20	13748	77.0	16755	93.9	11584	64.9
20-40	3687	20.7	956	5.4	3418	19.2
40-60	388	2.2	117	0.7	2437	13.7
60-80	17	0.1	13	0.1	383	2.1
80-100	3	0.0	2	0.0	10	0.1
All	17843	100.0	17843	100.0	17832	99.9
No population					11	0.1
Total villages					17843	100.0

Table 3
Work participation rates in different size class of villages

Work participation rate	I	II	III	IV	V	VI	VII	VIII	All
	0-100	101-500	501-1000	1001-2000	2001-3000	3001-5000	5001-10000	10001 and above	
Number of villages	249	2236	3889	5563	2756	2023	961	166	17843
All workers	48.22	49.35	48.09	46.61	46.16	43.90	44.31	42.07	45.22
All workers - Male	59.66	58.78	58.40	57.72	57.47	56.60	55.39	56.5	57.24
All workers - Female	35.67	39.49	37.32	34.99	34.33	30.51	32.73	26.66	32.57
Main workers - All	35.20	35.90	35.45	35.29	34.94	33.77	32.82	32.86	34.43
Main workers - Male	52.00	51.64	51.87	51.90	51.63	51.06	49.38	50.96	51.45
Main workers - Female	16.77	19.44	18.30	17.93	17.48	15.54	15.52	13.52	16.53
Marginal Workers - All	10.74	16.13	15.67	14.81	14.51	12.97	12.94	10.03	13.51
Marginal Workers - Male	18.35	26.43	25.91	24.46	24.03	21.76	21.77	17.24	22.52
Marginal Workers - Female	2.38	5.36	4.97	4.72	4.55	3.70	3.72	2.34	4.04
Main workers – Cultivator - All	9.51	13.39	13.35	13.22	12.81	11.62	10.97	10.85	12.32
Main workers – Cultivator - Male	12.63	16.71	16.94	16.68	16.27	15.30	14.13	14.65	15.89
Main workers – Cultivator - Female	6.08	9.91	9.60	9.61	9.18	7.72	7.66	6.78	8.57
Main workers – Agricultural labourers - All	0.15	0.21	0.28	0.29	0.30	0.37	0.43	0.39	0.32
Main workers – Agricultural labourers - Male	0.12	0.29	0.36	0.42	0.44	0.55	0.64	0.6	0.48
Main workers – Agricultural labourers - Female	0.18	0.13	0.20	0.15	0.16	0.18	0.21	0.16	0.16
Main workers – Household industry workers - All	14.81	6.17	6.15	6.97	7.32	8.82	8.48	11.59	8.27
Main workers – Household industry workers - Male	20.90	8.22	8.67	10.34	10.89	13.45	12.83	18.47	12.55

Work participation rate	I	II	III	IV	V	VI	VII	VIII	All
	0-100	101-500	501-1000	1001-2000	2001-3000	3001-5000	5001-10000	10001 and above	
Main workers – Household industry workers - Female	8.12	4.03	3.53	3.45	3.60	3.94	3.93	4.24	3.76
Main workers – Others - All	13.02	13.45	12.64	11.32	11.22	10.13	11.49	9.21	10.79
Main workers – Others - Male	7.66	7.14	6.53	5.82	5.84	5.54	6.01	5.54	5.8
Main workers – Others - Female	18.89	20.05	19.02	17.06	16.85	14.97	17.21	13.13	16.04
Marginal workers – Cultivator - All	2.08	2.26	2.37	2.36	2.29	1.84	2.36	1.38	2.05
Marginal workers – Cultivator - Male	0.99	1.05	1.04	1.01	0.98	0.85	1.32	0.75	0.93
Marginal workers – Cultivator - Female	3.27	3.53	3.77	3.77	3.65	2.87	3.44	2.04	3.23
Marginal workers – Agricultural labourers - All	7.27	8.85	8.24	7.07	6.93	6.12	6.85	5.53	6.66
Marginal workers – Agricultural labourers - Male	4.50	4.79	4.36	3.60	3.58	3.17	3.34	3	3.47
Marginal workers – Agricultural labourers - Female	10.31	13.09	12.29	10.71	10.43	9.23	10.51	8.24	10.02
Marginal workers – Household industry workers - All	0.05	0.12	0.16	0.13	0.14	0.15	0.12	0.16	0.15
Marginal workers – Household industry workers - Male	0.06	0.07	0.08	0.09	0.09	0.10	0.08	0.1	0.09
Marginal workers – Household industry workers - Female	0.03	0.18	0.25	0.17	0.2	0.2	0.17	0.22	0.2
Marginal workers – Others - All	3.62	2.21	1.86	1.75	1.86	2.02	2.16	2.15	1.93
Marginal workers – Others - Male	2.11	1.22	1.05	1.13	1.19	1.42	1.27	1.7	1.31
Marginal workers – Others - Female	5.28	3.25	2.71	2.41	2.56	2.66	3.08	2.63	2.59