Female Work Participation in Villages of India

Chaurasia, Aalok Ranjan

MLC Foundation and 'Shyam' Institute

20 January 2018

Female Work Participation in Villages of India

Aalok Ranjan Chaurasia

MLC Foundation
'Shyam' Institute

January 2018

Female Work Participation in Villages of India

Abstract

This paper attempts to identify patterns of female work participation in more than 578 thousand villages of India using the data mining approach. The analysis is based on an index of participation that has been developed for the purpose and takes into the consideration both the extent and the intensity of participation in productive activities. The analysis reveals that Indian villages can be grouped into 10 clusters with different level of female participation and with distinct village characteristics and there are distinct regional patterns. An interesting finding of the analysis is that participation of females in productive activities at the village level is relatively lower in villages higher level of female education as compared to villages with lower levels of female education. It appears that appropriate opportunities of participation for educate females are not available in the villages of India. Creating these opportunities at the village level is necessary not only for the transformation of village economy but also for women's empowerment.

Key Words

India, Villages, Females, Participation index, Data mining

Female Work Participation in Villages of India

Introduction

The ability to earn an income by participating in the social and economic productive activities is widely regarded as an important dimension of women's empowerment, although participation in social and economic productive activities, by itself, may not be regarded as a sufficient condition to ensure an increase in the bargaining power of women and a substantial decision-making role for women within the family and the society (Mencher and Saradamoni,1982; Bardhan, 1985; Nagaraj, 1989; Bennet, 1992; Sinha, 2005). In any case, the working status of women as a major factor in their economic flexibility has been considered as an indicator of their overall well-being status in the society (Mammen and Paxon, 2000). Female participation in productive activities has also been widely recognised as a driver of economic growth and, therefore, women's work participation rates indicate the potential for economic growth (Verick, 2014).

Participation of females in the social and economic production system is, however, influenced by a host of economic and social factors that interact in a complex fashion at the level of the household and the society. Global evidence suggests that the key factors influencing women's participation in productive activities include educational attainment, age at the entry into marital union, level of fertility, degree of urbanisation and social norms that determine the role of women in the family and the society. In India, it is argued that the increase in the educational enrolment of young women, lack of participation opportunities, household income, etc. are some of the factors that influence female participation in productive activities (Chaudhary and Verick, 2014; Kapsos et al, 2014). In addition, social norms that decide women's role in the public domain continue to affect outcomes. Variation in female work participation rates across states of India has been found to be influenced by a wide gamut of demographic, social, household and regional factors, although none of these factors are found to be mutually exclusive and can independently explain the variation in female work participation rates across Indian states (Jose, 1989; Sinha, 2005). Previous studies also suggest that there is no universal set of factors that explains the variation in female work participation rate across Indian states (Sundaram, 1989; Dholakia and Dholakia, 1978; Gulati, 1975; Nath, 1970). This implies that any analysis of participation of females in productive activities within broad socio-economic and cultural perspective such as the country as a whole or state or even district and sub-district does not appear to be the right approach of understanding the dynamics of female work participation. Rather, it appears to be more appropriate to analyse female work participation at the grassroots level and to identify the factors and conditions that influence this participation. There is however little attempt in this direction. Analysis of female work participation in India has generally been carried out at national and state levels (Sanghi et al, 2015; Vinoj, 2013; Rangarajan et al, 2011; Chaudhary, 2011;

Chandrasekhar and Ghosh, 2011). Village level analyses of female work participation are rare in India. Sinha (2005) has carried out an analysis of female participation in productive activities in villages of four districts of West Bengal whereas Rogers (2012) has used village level surveys to analyse labour force participation in 36 villages of Bihar. These studies have highlighted variation in female work participation across villages because of a number of village level factors. There has, however been little attempt to analyse variation in female work participation across the villages of the country. There has also been little attempt to identify village level factors that contribute to variation in female work participation across villages.

The above considerations constitute the rationale for the present analysis which aims at analysing the female work participation across the villages of India. We measure, in this paper, female work participation in more than 578 thousand villages of the country and explore distinct patterns of female work participation across villages on the basis of a set of village level characteristics following the data mining approach which is the process of discovering patterns in big data so as to extract information and transform the information into an understandable structure from the perspective of policy level analysis and programme level interventions (Hastie et al, 2013).

Participation in the social and economic productive activities, essentially, has two dimensions - the dimension of engagement and the dimensions of the duration of the engagement. The dimension of engagement reflects the extensiveness of participation while the dimension of duration of engagement indicates the intensity of participation. The two dimensions, in combination, determine the level of participation (Heckman, 1993). There is however no study in India, to the best of our knowledge, that takes into account both the dimensions of participation in studying female work participation. Female work participation has generally been analysed in terms of female labour force participation rate or the female work participation rate which is defined as the proportion of females engaged in any social and economic productive activity. Labour force participation rate and work participation rate are actually indicators of the extensiveness of participation. They do not take into account the intensity of participation in social and economic productive activities.

In this paper, we develop an index of participation in social and economic productive activities that takes into account both the dimensions of participation in productive activities - the dimension of extensiveness and the dimension of intensiveness - and use this index to measure and analyse female work participation across village of India. The analysis leads to the classification or segmentation of villages of the country into mutually exclusive groups or clusters of villages having distinct village level characteristics that are associated with different levels of female participation in village level social and economic production system. The analysis reveals that female participation at the village level is influenced by the defining characteristics of the village. The analysis also reveals that the composition of the female workforce at the village level varies across different groups or clusters of villages having distinct village level characteristics.

The paper is organised as follows. The next section of the paper describes the data used in the present analysis. The analysis is based on the primary census abstract of the 2011 population census which is the only source of data related to participation of the people in productive activities at the village level in India. Section three of the paper outlines the methods adopted for the analysis including a description of the index of participation that has been used to measure participation in the social and economic production system. Section four of the paper presents and discusses results of the analysis while the last section of the paper summarises main findings of the analysis and discusses their implications in the context of improving the participation of females in the productive activities at the village level.

The Index of Participation

The level of participation in productive activities in a village can be measured in terms of the total duration of engagement of the people of the village in productive activities within a reference period which may be a day or a month or a year or even a lifetime (Blundell at al, 2011). Assuming the reference period to be a year, the level of participation in the productive activities is then the product of the number of people engaged in productive activities during the year and the average number of days of engagement per person per year. If the number of people engaged in the productive activities in a village is L and the average number of days of engagement per person per year in the village is A, then the total duration of engagement, D, of the people of the village in productive activities in a year is

$$D = L * A$$

The above conceptualisation suggests that if E is the proportion of the population of the village who is engaged in productive activities in a year and I is the ratio of the average number of days of engagement of a person to the total number of days in the year, then the index of participation in productive activities, P, may be defined as

$$P = E * I$$

It is obvious that the higher is the value of the index P, the higher is the level of participation in productive activities. Here, the index E measures the extensiveness of the participation in productive activities and is nothing but the conventional work participation rate. The higher is the number of people engaged in productive activities, the higher is the value of E. On the other hand, the index I measures the intensity of participation in productive activities. The higher is the average duration of engagement of a person in productive activities in a year, the higher is the value of I. It is also obvious by definition that both E and I vary between 0 and 1 and, therefore, it is also obvious that the index of participation, P, defined as the product of the measure of the extensiveness and the measure of the intensiveness of participation measured in terms of the average duration of participation per person in a year also varies between 0 and 1.

The index of participation in productive activities, *P*, defined in the above manner, may be calculated for specific occupational categories also. If *Ej* is the extensiveness of participation and *Ij* is the intensity of participation in the in productive activity *j*, then the activity specific index of participation, *Pj*, may be defined as

$$P_j = E_j * I_j$$

The index of participation for all productive activities may then be defined as

$$P = \sum_j P_j = \sum_j E_j * I_j$$

The difference in the index of participation between two administrative units may then be decomposed, following Kitagawa (1955) in the following manner:

$$\begin{aligned}
P_2 - P_1 &= \partial P \\
&= \sum_j E_{j2} * I_{j2} - \sum_j E_{j1} * I_{J1} \\
&= \sum_j E_{j2} * I_{j2} - E_{j1} * I_{J1} \\
&= \sum_j (E_{j2} - E_{j1}) * \bar{I}_j + \bar{E}_j * (I_{j2} - I_{j1}) \\
&= \sum_j (E_{j2} - E_{j1}) * \bar{I}_j + \sum_j \bar{E}_j * (I_{j2} - I_{j1}) \\
&= \sum_j \nabla E_j + \sum_j \nabla I_j \\
&= \nabla E + \nabla I
\end{aligned}$$

where

$$\bar{I}_j = (I_{j2} + I_{j1}) / 2, etc.$$

We use the above analytical formulation measure and analyse the participation of females in social and economic productive activities at the village level of India. For every village of the country, we estimate the index of participation for different work categories by calculating the category-specific index of extensiveness of participation and category-specific index of intensiveness of participation of females. The category-specific index of extensive of participation and the index of intensiveness of participation have then been combined to obtain the index of female participation in the social and economic productive activities in the village. It is obvious that female participation in the village social and economic production system depends upon the opportunities of participation available in the local level social and economic production system.

Data

The analysis is based on the primary census abstract of the 2011 population census (PCA 2011) which provides data on the work status of the population for every village of the country. Work, in India's 2011 population census, is defined as participation in any economically productive activity with or without compensation, wages or profit (Government of India, 2011). Participation may be physical and/or mental. Work, according to the 2011 population census, involves not only the actual work but also supervision and direction. Part time help or unpaid work on farm, family enterprise or in any other economic activity has also been classified as work. People who are engaged in cultivation or milk production even solely for domestic consumption are also classified as workers. Workers enumerated at the population census are classified into three categories on the basis of the duration they worked during the year prior to the census - workers who worked for at least 6 months; workers who worked for 3-6 months; and workers who worked for less than 3 months. during the year prior to the census. Workers are further classified into one of the four occupational categories - cultivators; agricultural labourers; household industry workers and other workers. A worker is classified as cultivator if she or he is engaged in cultivation of his or her own land or land owned by other individuals or institutions including government for payment in money, kind or share. Cultivation includes effective supervision or direction in cultivation. A person who has given out her/his land to another person or persons or institution(s) for cultivation for money, kind or share of crop and who does not even supervise or direct cultivation of land, is not treated as cultivator. Similarly, a person working on another person's land for wages in cash or kind or a combination of both is not treated as cultivator. Cultivation involves ploughing, sowing, harvesting and production of cereals and millet crops such as wheat, paddy, jowar, bajra, ragi, etc., and other crops such as sugarcane, tobacco, ground-nuts, tapioca, etc., and pulses, raw jute and kindred fibre crop, cotton, cinchona and other medicinal plants, fruit growing, vegetable growing or keeping orchards or groves, etc. Cultivation does not include the following plantation crops - tea, coffee, rubber, coconut and betel-nuts (areca).

On the other hand, a person who works on another person's land for wages in money or kind or share is classified as agricultural labourer. She or he has no risk in the cultivation, but merely works on another person's land for wages. An agricultural labourer has no right of lease or contract on the land on which she/he works. Similarly, a worker is classified as household industry worker if she or he is engaged in a household industry which is an industry conducted by one or more members of the household at home or within the village in rural areas and only within the precincts of the house where the individual lives in urban areas. The household industry is not run on the scale of a registered factory. In the urban areas, even if household members run an industry by themselves but at a place away from the precincts of their home, it is not considered as a household industry. Household industry relates to production, processing, servicing, repairing or making and selling (but not merely selling) of goods. It does not include professions such as a Pleader, Doctor, Musician, Dancer, Waterman, Astrologer, Dhobi,

Barber, etc., or merely trade or business. Lastly, all workers who are not classified as either cultivators or agricultural labourers or household industry workers are classified as other workers. They include government servants, municipal employees, teachers, factory workers, plantation workers, workers engaged in trade; commerce; business; transport; banking; mining; construction; political or social work, priests, entertainment, artists, etc.

The concept of the village adopted in the Indian population census is different from the commonly used concept of the village as a human settlement which is larger than a hamlet but smaller than a town. A hamlet has a tiny population less than 100 (Doxiadis, 1968). During the census, the urban areas are identified first on the basis of clearly laid down definition of a standard urban area and the population living in these urban areas is classified as the urban population. Population not residing in the urban areas, on the other hand, is classified as the rural population which is then organised into administrative areas following the administrative boundaries of revenue villages and these administrative areas are termed as villages in the census parlance. A village, defined during the population census in India, is, therefore, an administrative unit with well-defined administrative boundaries and the population of the village so defined is the number of persons enumerated within the administrative boundaries of the village at the time of the census. This approach of defining a village during the census means that a village may have one or more than one human settlement within its administrative boundaries or it may have no human settlement at all in which case, the population of the village is zero. If there are more than one human settlements within the administrative boundaries of a village then PCA 2011 provides data of all human settlements combined and not data pertaining to separate human settlements within the village. This approach of defining a village pays no attention to the permanent or temporary nature of human settlements as the enumeration is carried out on the *de-facto* basis and not on the *de-jure* basis. There is always a possibility that human settlements within the boundaries of a village are permanent or temporary or both.

According to PCA 2011, there were 640949 villages in the country at the time of 2011 population census. There were 43330 or 6.8 per cent villages where no population was enumerated at the time of 2011 population census so that the number of inhabited villages in 2011 was 597619 or 93.2 per cent of the total villages identified. The population of these villages varied from just 1 to 66062 persons. In 18841 or 2.9 per cent villages, the number of households was less than 10. Moreover, there were 23 inhabited villages where there was no female population. We have excluded all these villages from the present analysis so that the present analysis is restricted to 578755 villages or the lowest level administrative units in the rural areas of the country as identified at the time of the 2011 population census. These villages account for 90 per cent of the villages of the country listed at the time of 2011 population census and the population living in these villages account for 99.9 per cent of the total rural population of the country - population living in those areas which have not been classified as urban at the time of the 2011 population census.

Methods

Estimating the index of participation, *P*, for the villages of the country requires estimation of the index of extensiveness of participation, *E*, and the index of intensiveness of participation in the village. We have measured the index of extensiveness of participation in terms of the ratio of the workers to the total population in the village. A more appropriate indicator to measure the extensiveness of participation would have been the ratio of workers aged 15-50 years to the population aged 15-59 years. However, PCA 2011 does not provide data pertaining to population and workers in the villages by age. Although, the ratio of workers to total population is a crude measure of the extensiveness of participation, yet, it gives an idea about how many people in the village are engaged in productive activities - the higher is this ratio, the higher is extensiveness of participation.

On the other hand, the index of intensiveness of participation, *I*, is estimated as the ratio of average annual days of work per person per year divided by the potential number of days of work for a person during one year. For estimating the average annual number of days of work per person, we assumed that frequencies of each work interval are centred at the mid vale of the interval. Thus, we assumed that workers who worked for at least 6 months during the year prior to the census actually worked for 270 days in a year - 270 is approximately the mid value of the interval 180-365 days. Similarly, we assumed that workers who worked for 3-6 months in the year prior to the census actually worked for 135 days in a year and workers who worked for less than 3 months actually worked for 45 days in the year. Under these assumptions, the average annual days of work in a year per person (*adw*) is calculated as

$$adw = \frac{(270 * W_M + 135 * W_G + 45 * W_L)}{W}$$

where W_M is the number of workers who worked for at least 6 months in the year, W_G is the number of workers who worked for 3-6 months in the year and W_L is the number of workers who worked for less than 3 months in the year. Once, *adw* is estimated, the index of intensiveness of participation, *I*, is than calculated as $I = adw/270$. Here, it is assumed that the potential days of work available in the village is 270 days.

It is logical to assume that the index of participation, *I*, is influenced by a host of village specific characteristics. These characteristics include but are not limited to 1) population of the village; 2) level of education in the village; 3) level of fertility in the village; 4) gender composition of the village population; and 5) social class structure of the village population. We have measured the level of education in the village in terms of the effective literacy rate which is defined as the proportion of population aged 7 years and above who can read and write with understanding. Similarly, the level of fertility has been surrogated by the ratio of children aged 0-6 years to females aged 7 years and above. This ratio is very similar to the familiar child-woman ratio which is widely used as a crude indicator of fertility (Shryock and Siegel, 1980). Obviously, the higher is this ratio, the higher is the level of fertility.

In order to analyse how village characteristics influence the participation in productive activities in the village, we have followed the classification modelling approach. This approach classifies or segments villages into different groups or cluster of villages in such a manner that the within-group or within cluster homogeneity with respect to the index of participation is the highest. The classification modelling approach is different from the regression-based approach that is commonly used for analysing marginal effects of the defining characteristics of the village on the index of participation (Chaurasia, 2012). Unlike the regression-based approach, there is no restriction or limitation on the structure of the independent variables or the defining characteristics of the village which are used as explanatory variables in the classification modelling exercise. In the most general terms, the classification or the segmentation emanating from the classification modelling exercise is based on a set of *if-then* logical conditions that permit splitting or classifying or segmenting villages into mutually exclusive groups or classes of villages.

We have used the Decision Tree procedure to create a tree-based classification or segmentation model which classifies villages into groups or clusters of villages by partitioning villages into smaller groups or classes so that villages within a group are as homogenous as possible. The procedure can be used for many purposes and we use the procedure here for the segmentation of villages (IBM Corporation, 2012). Within group or cluster homogeneity is measured in terms of entropy. The lower is the entropy, the higher is the homogeneity. There are many tree-growing methods which can be used for classification modelling. We have used the classification and regression tree (CRT) method (Breiman et al, 1984) which is a nonparametric recursive partitioning method. CRT splits villages into groups or segments that are as homogenous as possible with respect to the dependent variable - index of participation. A terminal node in which all villages have the same value for the dependent variable is called a homogeneous, "pure" node. The extent to which a node does not represent a homogenous subset of villages is an indication of impurity. There are different impurity measures available. In case of continuous or scale variables, impurity is measured in terms of least-squared deviation which is computed as the within-node variance adjusted for frequency weights or influence variables as the case may be (IBM Corporation, 2012).The tree-growing process is continued until either the pure node is reached or the prescribed stopping criterion is met (Ambalavanan et al, 2006; Lemon et al, 2003).

CART has a number of advantages as an exploratory data analysis procedure for the purpose of classification or segmentation which is the primary objective of the present analysis. It makes no assumption about the distribution of the dependent variable or the independent variables used in the analysis. Moreover, explanatory or the independent variables used in CART can be a mix of categorical, interval, and continuous or scale variables. Another advantage of CART is that results of the analysis are not at all affected by the quality of data such as presence of outliers, collinearity among explanatory variables, heteroscedasticity, or distributional error structures that normally affect parametric procedures.

Female Participation in Villages of India

On the basis of the data available through the 2011 population census, we have estimated the index of female participation in productive activities, P_f, for 578755 villages included in the present analysis. P_f is found to vary widely across villages of the country (Table 1). There are 2954 villages in the country where P_f is estimated to be zero whereas there are 20 villages where P_f is estimated to be 1 at the time of 2011 population census. The distribution of villages by P_f has been found to be positively skewed which means that in majority of the villages, P_f is less than the average. In nearly two-third villages of the country, P_f is estimated to be less than 0.30 and in almost 38 per cent villages, it is estimated to be less than 0.15. By contrast, there are only about 3 per cent villages where P_f is estimated to be 0.60 and more. The weighted average of P_f across 578755 villages is estimated to be 0.224 whereas the unweighted average is estimated to be 0.250. The weighted average takes into account the size of the population of the village which varies widely across villages. The fact that the weighted average of P_f is lower than its unweighted average implies that P_f is relatively lower in larger villages than in smaller villages of the country.

Table 1 also summarises the distribution of villages in terms of the index of female participation by work categories - cultivation (P_{fc}), agricultural labour (P_{fa}), household industry work (P_{fh}) and in other productive activities (P_{fo}). The table indicates that the index of female participation in household industry, P_{fh}, is very low. There are only about 1 per cent villages where P_{fh} is at least 0.15. Similarly, there are less than 5 per cent villages where P_{fo} is at least 0.15. By contrast, the index of female participation in cultivation, P_{fc} and in agricultural labour, P_{fa} is estimated to be 0.15 and more in about 23 per cent and 29 per cent villages of the country respectively. At the same time, there are 37 villages in the country where the index of female participation in household level productive activities, P_{fh}, is estimated to be very high - 0.60 and more. Similarly, in 272 villages of the country, the index of female participation in productive activities other than cultivation, agricultural labour and household industrial activity, P_{fh}, is estimated to be very high.

The index of female participation in productive activities may also be viewed as a reflection of the participation opportunities available for females in the village level social and economic production system and it can be assumed that wider are the opportunities for participation of females available in the village level social and economic production system, the higher is the female participation in the productive activities. In this context, table 1 suggests that the opportunities for female participation in village level productive activities appear to be seriously limited in nearly all the villages of the country and whatever opportunities are available they are virtually confined to the agriculture sector, particularly, agricultural labour. There appears to be very limited opportunities of productive participation of females in sectors other than agriculture at the village level and this appears to be an important reason behind low to very low participation of females in village social and economic production system.

The index of female participation, P_f, is determined by the index of extensiveness of female participation, E_f, and the index of intensiveness of female participation, I_f. Table 2 presents the weighted average of E_f and I_f by occupational category across villages of the country. The E_f and I_f for all occupational categories combined is the sum of occupational category specific E_f and I_f. It may be seen from the table that the variation in E_f across occupational categories is more marked than the variation in I_f across occupational categories. Moreover, E_f is the highest in the occupational category agricultural labour but I_f in this occupational category is very low. On the other hand, both E_f and I_f are the lowest in the occupational category household industry.

The difference in the index of female participation P_f between two occupational categories is due to the difference in both the index of extensiveness E_f and the index of intensiveness I_f. It is possible to decompose the difference in P_f between two occupational categories into the difference between E_f and the different between I_f. We have carried out this decomposition to find out how much of the difference in the index of female participation as agricultural labour (P_{fa}) and as household industry worker (P_{fh}) is attributed to the difference in the index of extensiveness and to the difference in the index of intensiveness. This exercise suggests that almost all the difference between P_{fa} and P_{fh} is attributed to the difference in the index of extensiveness. The contribution of the difference between the intensiveness of female participation as agricultural labour and as household industry worker and the joint contribution of the difference in the extensiveness and the intensiveness of participation is at best marginal. This observation suggests that the very low index of female participation in the household industry in the villages of the country is largely due to very limited opportunities of female participation in household level productive activities. This means that if female participation in the household industry sector is to be increased, then it is imperative that opportunities for female participation in household level productive activities at the village level must be expanded. The intensiveness of female participation is not a major issue in improving female participation in the household productive activities at the village level. It is also clear from table 2 that the same argument applies to increasing the female participation in other sectors of the village social and economic production system which is also, at present, is very low.

Classification of Villages

Data available through the 2011 population census clearly shows that female participation in productive activities varies widely across the 578755 villages of India. It is also well known that Indian villages very widely in terms of selected defining characteristics of the village population. This means that variation in the index of female participation in village level productive activities needs to be analysed in terms of the variation in a set of defining characteristics of the village population. We have carried out classification modelling exercise to examine this association. The classification modelling exercise essentially involved building the classification tree through CRT methodology.

The dependent variable in the classification modelling exercise is the index of female participation P_f which is estimated for 578755 villages of the country. The independent or classification variables used for the purpose of modelling include: 1) total population of the village; 2) population sex ratio measured in terms of the ratio of the number of males aged 7 years and above to the number of females aged 7 years and above in the village; 3) proportion of Scheduled Tribes females to total female in the villages; 4) effective female literacy rate in the village measured in terms of the proportion of females aged 7 years and above in the village who can read and write with understanding; and 5) level of fertility in the village measured in terms of the ratio of the population aged 0-6 years to females aged 7 years and above.

Results of classification modelling exercise are presented in table 3 and the classification tree is depicted in figure 2. The first split of 578755 villages of the country included in the present analysis is on the proportion of Scheduled Tribes females to total females in the village - 307100 (53.1 per cent) villages where there was virtually no Scheduled Tribes population at the 2011 population census (Node 1) and 271655 (46.9 per cent) villages where Scheduled Tribes were present (Node 2). The unweighted average of P_f across villages of Node 1 is substantially lower than that in villages of Node 2. Next, villages of Node 1 are further split on the level of fertility in the village - 127103 (22.0 per cent) villages where the proportion of children aged 0-6 years to females aged 7 years and above is less than or equal to 0.316 (Node 3) and 179997 (31.1 per cent) villages where this ratio is more than 0.316 (Node 4). The unweighted average of the P_f across villages of Node 3 is 0.242 compared to 0.178 across villages of Node 4. On the other hand, villages of Node 2 are further split into 181084 (31.1 per cent) villages where the proportion of Scheduled Tribes females to total females in the village ranged between 0.003-0.638 (Node 5) and 90571 (15.6) villages where the proportion of Scheduled Tribes females to total females in the village is more than 0.638 (Node 6). The unweighted average of the P_f across villages of Node 5 is 0.281 compared to 0.339 across villages of Node 6.

At the third level of classification, villages of Node 3 are further split into 46637 (8.1 per cent) villages where effective female literacy is less than or equal to 0.628 (Node 7) and 80466 (13.9 per cent) villages where the effective female literacy is more than 0.628 (Node 8). The unweighted average of P_f across villages of Node 7 is 0.295 compared to 0.212 across villages of Node 8. On the other hand, villages of Node 4 are split further on the population of the villages into 75467 (13.0 per cent) villages having population less than or equal to 842 (Node 9) and village having population more than 842 (Node 10). The unweighted average of P_f across villages of Node 9 is 0.208 compared to 0.157 across villages of Node 10. Similarly, villages of Node 5 are split into 58384 (10.1 per cent) villages having effective female literacy rate less than or equal to 51.3 per cent (Node 11) and 122700 (21.2 per cent) villages having effective female literacy rate of more than 51.3 per cent (Node 12). The unweighted average of P_f across villages of Node 11 is 0.308 compared to 0.268 across villages of Node 12. There is no further splitting of villages of Node 6 so that Node 6 is a terminal node.

At the last level of classification, villages of Node 7 are further split on the ratio of children 0-6 years to females aged 7 years and above into 16837 (2.9 per cent) villages where this ratio is less than or equal to 0.243 (Node 13) and 29800 (5.1 per cent) villages where this ratio ranges between 0.243-0.316 (Node 14). The unweighted average of P_f across villages of Node 13 is 0.352 compared to 0.283 across villages of Node 14. On the other hand, villages of Node 8 are split on the ratio of males aged 7 years and above to females 7 years and above into 31170 (5.4 per cent) villages with this ratio less than or equal to 1.001 (Node 15) and villages with this ratio greater than 1.001 (Node 16). The average of P_f across villages of Node 15 is 0.255 compared to 0.185 across villages of Node 16. There is no further split of villages of Node 9 and Node 10 whereas villages of Node 11 are split further on the ratio of children 0-6 years to females 7 years and above into 15384 (2.7 per cent) villages with this ratio less than or equal to 0.316 (Node 17) and 43000 (7.4 per cent) villages having this ratio more than 0.316 (Node 18). The unweighted average of P_f across villages of Node 17 is 0.382 compared to 0.281 across villages of Node 18. Finally, there is no further split of villages of Node 12.

As may be seen from figure 2, the classification exercise yields 10 terminal nodes which means that 578755 villages of the country can be grouped into 10 mutually exclusive Nodes or clusters of villages in the context of the index of female participation in productive activities at the village level. The defining characteristics of villages of different clusters are shown in table 3. The unweighted average of P_f of villages of cluster 10 is the lowest but that of cluster 17 is the highest amongst the 10 clusters. Other clusters where the unweighted average of P_f is relatively high are cluster 13 and cluster 6 whereas the unweighted average of P_f is relatively low in cluster 16 and cluster 9. It may also be observed from figure 2 and table 3 that within cluster distribution of villages in terms of P_f is generally positively skewed with the exception of cluster 13 and cluster 17 where the within cluster distribution of villages in terms of P_f. A positive skewness means that in majority of the villages within the cluster, P_f is less than the average P_f for the cluster but there are villages where P_f is higher than the cluster average. The degree of skewness in within cluster distribution of villages by P_f, however, varies across clusters being very high in cluster 16 and cluster 12 but very low in cluster 6 and cluster 18. In clusters 13 and 17, however, P_f is higher than the cluster average in majority of the villages in the cluster but there are villages in these clusters where P_f is smaller than the unweighted cluster average.

The patterns of female participation in productive activities in different clusters of villages are presented in table 4. The variation in the index of female participation in productive activities, P_f, across clusters is due to both variation in the index of extensiveness of participation, E_f, and variation in the index of intensiveness of participation, I_f. There are only three clusters where more than 40 per cent of females were found to be engaged in productive activities at the time of 2011 population census whereas in two clusters, this proportion was less than 25 per cent. On the other hand, there are only two clusters where the intensity of participation was more than 80 per cent whereas in two clusters, it was less than 70 per cent.

More specifically, the very high index of female participation, P_f, in cluster 17 is due to the highest index of the extensiveness of participation E_f and the second highest index of intensiveness of participation I_f. Similarly, P_f is the second highest in cluster 13 because E_f is the third highest in the cluster but I_f is the highest. By contrast, P_f is the lowest in cluster 10 because E_f is the lowest and I_f is the second lowest. Similarly, P_f is the second lowest in cluster 16 primarily because E_f is the lowest in this cluster, although I_f ranks sixth in this cluster. In cluster 9, P_f is the third lowest primarily because I_f is the lowest in this cluster, although E_f ranks seventh across clusters. Table 4 also reveals that below average E_f in clusters 12 and 15 is associated with the above average I_f in these clusters. There is however no cluster in which both E_f and I_f are the highest or the lowest across clusters.

Table 4 also presents patterns of female participation across clusters by different occupational categories. The index of female participation as cultivator, P_{fc} is the highest in cluster 6, not in cluster 17. However, the index of female participation as agricultural labourer and household industry workers is the highest in cluster 17. On the other hand, the index of female participation as other occupations is the highest in cluster 15. By contrast, female participation as cultivator, agricultural labour and other occupations is the lowest in cluster 10 but participation in household productive activities is the lowest in cluster 6. It is evident from table 4 that these variations in the index of female participation can be traced in variation in both the index of extensiveness and the index of intensiveness of participation in different occupations across clusters identified through classification modelling. It may also be observed from table 4 that very high index of extensiveness of female participation, P_f, in clusters 6, 13 and 17 is largely because of very high index of extensiveness of female participation as agricultural labourers. At the same time, very low index of extensiveness in female participation, E_f, in cluster 10 is primarily because of very low index of extensiveness of female participation as cultivators, P_{fc}, and agricultural labourers, P_{fa}. Similarly, very high index of intensiveness of female participation, I_f, in clusters 13 and 17 is largely because of very high index of intensiveness of female participation in all the four occupational categories. By comparison, the index of intensiveness of female participation in all the four occupational categories is generally very low in cluster 10. These observations suggest that opportunities for female participation in village level productive activities vary widely across the 10 clusters of villages identified through classification modelling.

The difference in the index of female participation, P_f between two clusters may be decomposed into the difference resulting from the difference in the index of intensiveness of female participation, E_f, and the difference resulting from the difference in the index of intensiveness of female participation, I_f. Results of this decomposition analysis are presented in table 5. For example, the difference between the index of female participation in cluster 17, the cluster with highest P_f and cluster 10, the cluster with lowest P_f is 0.261 in absolute terms. More than 78 per cent of this difference is attributed to the difference in the index of extensiveness of female participation, E_f, between the two clusters while the difference in the index of intensiveness of female participation, I_f,

accounts for about 22 per cent of the difference in P_f. More specifically, the difference in the index of extensiveness of female participation as agricultural labour, E_{fa} alone accounts for more than 56 per cent of the difference in the index of female participation, P_f, between cluster 17 and cluster 10. At the same time, the difference in the intensiveness of female participation as agricultural workers between the two clusters accounts for almost 13 per cent of the total difference. In other words, the very high participation of females in villages of cluster 17 is primarily due to very high index of female participation as agricultural labourers in cluster 17 as compared to cluster 10. In cluster 17, not only the index of extensiveness of female participation as agricultural worker is higher than that in cluster 10 but also the index of intensiveness of female participation as agricultural labourer is also higher.

Table 5 also shows that the difference in the index of female participation of different clusters from the index of female participation in cluster 10 may be attributed to the difference in female participation in different occupational categories. For example, the difference in the index of female participation of cluster 15 and cluster 16 from that of cluster 10 may be attributed largely to the difference in the index of female participation as other workers whereas this difference in clusters 6 and 9 may be attributed largely to the difference in the index of female participation as cultivators. In case of other clusters, this difference is attributed largely to the difference in the index of female participation as agricultural labourers. Even in case of cluster 6, the difference in the index of female participation as agricultural labourer is quite substantive. Interestingly, the difference in the index of female participation as household industry workers is not found to be substantial in any cluster.

It may also be seen from table 5 that the contribution of the difference in the index of extensiveness of female participation is substantially higher than that of the index of the intensiveness of female participation in all but one cluster. One reason for the relatively low contribution of the difference in the index of intensiveness of female participation to the index of female participation may be the way the index of intensiveness of female participation has been calculated because of data constraints. We have assumed in calculating the index of intensiveness in female participation that the engagement in a productive activity is either for 45 days in a year or 135 days in a year or 270 days in a year on average. The reason for this hard assumption is that PCA 2011 provides data about the number of females who worked for 1-90 days in a year; 90-180 days in a year and 180-365 days in a year. For the calculation of the index of intensiveness of female participation, the middle value of each interval has been taken. During the 2011 population census, data pertaining to the actual number of days of engagement in productive activities during the year prior to the census was not collected from the working population. Rather, every worker enumerated at the census was asked whether she or he worked for 6 month or more or for 3-6 months or for less than 3 months during the year prior to the census. The estimates of average duration of engagement in productive activities are therefore derived from the limited set of data. As such, variation in the index of intensiveness of female participation across villages is substantially limited.

Regional Patterns in Female Work Participation

Table 6 presents the distribution of villages of different states/Union Territories by different clusters. For the country as a whole, almost 68 per cent of the villages are confined to clusters 6, 9, 10 and 12 but villages of different states/Union Territories are largely confined to different clusters. More than 95 per cent villages of Kerala are confined to clusters 12 and 15 only but more than 95 per cent villages in Delhi and more than 87 per cent villages in Haryana are confined to clusters 10 and 16 only. Almost 80 per cent villages in Bihar and almost 75 per cent villages in Uttar Pradesh are confined to clusters 9 and 10. In the north-eastern states of the country, virtually all villages belong to cluster 6 only. In other states/Union Territories, villages are not confined to one or two clusters. For example, villages in Jammu and Kashmir are almost equally distributed across clusters 9, 10, 12 and 18. The same appears to be the case in Rajasthan where majority of villages are distributed almost equally across five clusters - 6, 9, 10, 12 and 18. In Andhra Pradesh, more than 87 per cent villages are distributed across five clusters of villages 6, 12, 13, 14 and 17. Similarly, almost 93 per cent villages in Tamil Nadu are distributed across only five clusters.

Viewed differently, more than 40 per cent villages of cluster 6 are located in - Jharkhand, Chhattisgarh and Madhya Pradesh. Almost 55 per cent villages of cluster 9 are located in Rajasthan, Uttar Pradesh and Bihar, with Uttar Pradesh alone accounting for almost one third villages of this cluster. Nearly two third villages of cluster 10 are located in Uttar Pradesh and Bihar with Uttar Pradesh accounting for almost 45 per cent villages of this cluster. Similarly, almost 54 per cent of villages of cluster 13 are located in Andhra Pradesh, Karnataka and Tamil Nadu and around 43 per cent villages of cluster 14 are located in Rajasthan, Uttar Pradesh and Andhra Pradesh. More than half of the villages of cluster 15 are located in Himachal Pradesh, Uttarakhand, Odisha and Tamil Nadu where as around 53 per cent villages of cluster 16 are located in only four states - Punjab, Uttar Pradesh, West Bengal and Odisha. More than three fourth villages of cluster 17, on the other hand, are located in only four states - Rajasthan, Odisha, Andhra Pradesh and Karnataka - with Andhra Pradesh, alone, accounting for more than 42 per cent villages of this cluster. Finally, more than half of the villages of cluster 18 are located in only three states - Rajasthan, Jharkhand and Madhya Pradesh.

Conclusions

The present analysis has attempted to measure participation in social and economic productive activities in more than 578 thousand villages of India on the basis of an index of female participation that takes into account both the extensiveness and the intensiveness of participation in productive activities. The analysis indicates that the level of female participation in the village economy in India remains far from satisfactory. There are very few villages in the country where female participation in productive activities may be termed as satisfactory. More importantly, there is only a nominal participation of

females in household level productive activities in all but a few villages and one probably reason may be that there are very limited opportunities of household level productive activities in the village economy in India. The level of female participation in the villages of the country is largely dependent upon the level of female participation as agricultural labour.

The analysis also suggests that the level of female participation in the village economy is influenced by the social class structure and the gender composition of the village population, the level of female education and the level of fertility in the village in addition to the size of the village population. The level of female participation is found to be associated positively with the proportion of Scheduled Tribes population in the village but negatively with the level of fertility, level of female education and the size of the village population. The level of female participation has also been found to be relatively lower in those villages where the gender balance is not in favour of females. When these five defining characteristics of the village are taken into consideration, more than 578 thousand villages of the country can be grouped into 10 mutually exclusive clusters, each having significantly different level of female participation. The 10 clusters of villages identified in the present analysis also have distinct regional patterns with villages of some clusters confined largely to one or two states/Union Territories only.

A revealing finding of the present analysis is that the level of female participation in village economy is negatively associated with the level of female education in the village as measured by the female literacy rate. This negative association contradicts the widely held argument that female education increases opportunities for the participation of women in productive activities. It appears that opportunities of participation of educated females in the economy of the villages of the country are very limited. If the index of extensiveness of female participation is any indication, then work, in the villages of the country, is largely confined to the agriculture sector and that too in the form of agricultural labour. It is argued that females prefer women-centric work which can be discharged from the household or within the household premises in the nature of self-employment (Sanghi et al, 2015). However, such opportunities of participation appear largely absent in the villages of India according to the data available through the 2011 population census. The village economy in India remains largely traditional with little sign of transition despite substantial expansion and growth of country's economy in the recent past. It appears that the growth and expansion of the economy of the country have largely been irrelevant to Indian villages where almost 70 per cent of India's population lives. This also means that the economic growth and associated development in India appears to have contributed little to rural women by expanding opportunities for the participation of females in the village level productive activities.

References

Ambalavanan N, Baibergenova A, Carlo WA, Saigal S, Schmidt B, Thorpe KE (2006) Early prediction of poor outcome in extremely low birth weight infants by classification tree analysis. The Journal of Pediatrics 148(4):438.e1-444.e1.

Bardhan K (1985) Women's work, welfare and status: Forces of tradition and change in India. *Economic and Political Weekly* 22(50, 51 &52).

Bennet L (1992) Women, poverty and productivity in India. *Seminar Paper No. 43*, Economic Development Institute. Washington DC, The World Bank.

Blundell R, Bozio A, Laroque G (2011) Extensive and intensive margins of labour supply: Working hours in the US, UK and France. Bonn, The Institute for the Study of Labour. Discussion Paper No. 6051.

Breiman L, Friedman J, Olshen RA, Stone CJ (1984) *Classification and Regression Trees*. Wadsworth, 1984.

Chandrasekhar CP, Ghosh J (2011) Latest employment trends from NSSO. Business Line, July 12.

Chaudhary R, Verick S (2014) Female labour force participation in India and beyond. New Delhi, International Labour Organization. ILO DWT for South Asia and ILO Country Office for India. ILO Asia-Pacific Working Paper Series.

Chaudhary S (2011) Employment data in India. What does latest data show? *Economic and Political Weekly* 46(32).

Chaurasia AR (2012) Contraceptive use in India: A data mining approach. *International Journal of Population Research*. Volume 2014, Article ID 821436, 11 pages. http://dx.doi.org/10.1155/2014/821436

Das S, Jain-Chandra S, Kochhar K, Kumar N (2015) Women workers in India. Why so few among so many. International Monetary Fund. IMF Working Paper WP/15/55.

Dholakia BH, Dholakia RH (1978) Inter-state variation in female labour force participation rates in India. *The Indian Journal of Labor Economics* 20(4).

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

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

Gulati L (1975) Female work participation: A study of inter-state differences. *Economic and Political Weekly* 10 (1 & 2).

Heckman J (1993) What has been learned about labor supply in the past twenty years? *American Economic Review* 83(2):116-121.

Hastie T, Tibshirani R, Friedman J (2008) *The Elements of Statistical Learning. Data Mining, Inference and Prediction*. Springer.

IBM Corporation (2012) *IBM SPSS Decision Trees 21*. IBM Corporation.

Jose AV (1989) *Limited Options: Women Workers in Rural India*, Asian Regional Team for Employment Promotion and World Employment Programme, International Labour Office.

Kapsos S, Silberman A, Bourmpoula E (2014) Why is female labour force participation declining so sharply in India? Geneva, International Labour Office. ILO Research Paper No. 10.

Kitagawa EM (1955) Components of a difference between two rates. *Journal of American Statistical Association* 50(272):1168-1194.

Lemon SC, Roy J, Clark MA, Friedmann 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):172-181.

Mammen K, Paxon C (2000). Women's work and economic development. *The Journal of Economic Perspectives* 14(4):141–164.

Mencher J, Saradamoni P (1982) Muddy feet, dirty hands: Rice production and female agricultural labour. *Economic and Political Weekly,* 17(52).

Nagaraj K (1989) Female workers in rural Tamil Nadu: A preliminary study. In AV Jose (ed) *Limited Options: Women Workers in Rural India*, Asian Regional Team for Employment Promotion and World Employment Programme, International Labour Office.

Nath K (1970) Female work participation and economic development. *Economic and Political Weekly* 15(31).

Rangarajan C, Kaul PI, Seema (2011) Where is the missing labour force? *Economic and Political Weekly* 46(39):68-72.

Rogers J (2012) Labour force participation in rural Bihar: A thirty-year perspective based on village surveys. New Delhi, Institute for Human Development. WP04/2012.

Sanghi S, Sirja A, Vijay SS (2015) Decline in rural female labour force participation in India. A relook into causes. *Vikalpa. The Journal of Decision Makers* 40(3):255-268.

Shryock SH, Siegel JS, Larmon A (1980) *The Methods and Materials of Demography*. Vol I. United States Bureau of Census.

Sinha S (2005) Female work participation rates in rural West Bengal: A village level analysis. *Indian Journal of Labour Economics* 48(3):563-577.

Sundaram K (1989) Inter-state variations in participation rates of women in India: An analysis. In AV Jose (ed) *Limited Options: Women Workers in Rural India*, Asian Regional Team for Employment Promotion and World Employment Programme, International Labour Office.

Verick S (2014) Women's labour force participation rate in India. Why is it so low? New Delhi, ILO DWT for South Asia and Country Office for India.

Vinoj A (2013) Missing labour force or 'defeminization' of labour force in India? Thiruvananthapuram, Centre for Development Studies. CDSWP 452.

Table 1
Distribution of villages by index of female participation, P_f, in different occupational categories in India, 2011

P_f		All	Cultivation	Agricultural labour	Household industry	Others
Very low	< 0.15	37.7	77.4	71.3	98.8	95.7
Low	0.15-0.30	26.7	12.3	19.8	0.9	3.3
Average	0.30-0.45	20.1	6.2	6.7	0.2	0.8
High	0.45-0.60	13.2	3.3	1.9	0.1	0.2
Very high	≥0.60	3.3	0.8	0.3	0.0	0.0
Weighted mean		0.224	0.068	0.104	0.011	0.041
Unweighted mean		0.250	0.096	0.107	0.009	0.038
Standard deviation		0.178	0.140	0.126	0.033	0.058
Median		0.228	0.026	0.056	0.001	0.019
IQR		0.284	0.123	0.164	0.005	0.033
Skewness		0.540	1.948	1.486	8.076	4.268

Source: Author's calculations

Table 2
Weighted average of the index of extensiveness and the index of intensiveness in villages of India by occupation category

Index	All	Cultivation	Agricultural labour	Household industry	Others
E_f	0.300	0.087	0.145	0.015	0.053
I_f	0.748	0.790	0.716	0.711	0.775
P_f	0.224	0.069	0.104	0.011	0.041

Source: Author's calculations

Table 3
Female work participation rate by defining characteristics of villages

Node/ Cluster number	Proportion of Scheduled Tribes females	Male/Female ratio in population aged at least 7 years	Ratio of children 0-6 years to females aged at least 7 years	Female effective literacy rate	Population	Distribution of the index of female participation (P_f) within the cluster			Number of villages in the cluster
						Unweighted average	Standard deviation	Skewness	
6	> 0.638					0.339	0.168	0.198	90571
9	≤ 0.003		> 0.316		≤ 842.5	0.208	0.181	0.771	75467
10	≤ 0.003		> 0.316		> 842.5	0.157	0.132	1.171	104530
12	0.003-0.638			> 0.513		0.268	0.165	0.423	122700
13	≤ 0.003		≤ 0.243	≤ 0.628		0.352	0.194	-0.077	16837
14	≤ 0.003		0.243-0.316	≤ 0.628		0.263	0.184	0.420	29800
15	≤ 0.003	≤ 1.001	≤ 0.316	> 0.628		0.255	0.187	0.572	31170
16	≤ 0.003	> 1.001	≤ 0.316	> 0.628		0.185	0.175	1.214	49296
17	0.003-0.638		≤ 0.316	≤ 0.513		0.382	0.162	-0.269	15384
18	0.003-0.638		> 0.316	≤ 0.513		0.281	0.156	0.274	43000
All						0.250	0.178		578755

Source: Author's calculations

Table 4
Effectiveness and intensiveness of female participation in different clusters

Index	Node/Cluster										
	6	9	10	12	13	14	15	16	17	18	All
						All workers					
E_f	0.470	0.290	0.213	0.318	0.439	0.332	0.286	0.221	0.487	0.370	0.300
I_f	0.712	0.686	0.690	0.786	0.848	0.770	0.786	0.765	0.839	0.706	0.748
P_f	0.335	0.199	0.147	0.250	0.372	0.256	0.225	0.169	0.409	0.261	0.224
						Cultivators					
E_{fc}	0.176	0.114	0.058	0.080	0.123	0.101	0.084	0.061	0.111	0.118	0.087
I_{fc}	0.797	0.732	0.734	0.825	0.887	0.806	0.754	0.784	0.905	0.768	0.790
P_{fc}	0.140	0.083	0.043	0.066	0.109	0.081	0.063	0.048	0.100	0.091	0.069
						Agricultural labourers					
E_{fa}	0.235	0.119	0.098	0.162	0.235	0.163	0.101	0.083	0.299	0.191	0.145
I_{fa}	0.653	0.634	0.647	0.763	0.831	0.749	0.785	0.740	0.815	0.669	0.716
P_{fa}	0.153	0.075	0.063	0.124	0.195	0.122	0.079	0.061	0.244	0.128	0.104
						Household industry workers					
E_{fh}	0.011	0.013	0.016	0.015	0.018	0.013	0.015	0.016	0.027	0.013	0.015
I_{fh}	0.619	0.650	0.679	0.735	0.840	0.730	0.791	0.698	0.860	0.651	0.711
P_{fh}	0.007	0.008	0.011	0.011	0.015	0.009	0.012	0.011	0.023	0.008	0.011

Index	Node/Cluster										
	6	9	10	12	13	14	15	16	17	18	All
						Other worker					
E_{fo}	0.048	0.044	0.041	0.062	0.062	0.056	0.086	0.062	0.052	0.048	0.053
I_{fo}	0.711	0.722	0.735	0.810	0.840	0.776	0.817	0.796	0.825	0.716	0.775
P_{fo}	0.034	0.032	0.030	0.050	0.052	0.043	0.070	0.049	0.043	0.034	0.041

Source: Author's calculations

Table 5

Decomposition of the difference of the index of female participation, P_f, in a cluster from the index of female participation in cluster 10, the cluster with the lowest index of female participation.

Cluster	∇P_f	Difference attributed to (Per cent)													
		∇E_f	∇I_f	Cultivation			Agricultural labour			Household industry work			Others		
				∇P_{fc}	∇E_{fc}	∇I_{fc}	∇P_{fa}	∇E_{fa}	∇I_{fa}	∇P_{fh}	∇E_{fh}	∇I_{fh}	∇P_{fo}	∇P_{fo}	∇P_{fo}
16	0.022	32.1	67.9	21.2	7.6	13.6	-8.0	-46.6	38.6	0.8	-0.6	1.4	86.0	71.7	14.3
9	0.052	105.2	-5.3	78.8	79.3	-0.5	22.9	25.8	-2.9	-4.4	-3.6	-0.8	2.6	3.7	-1.1
15	0.077	71.4	28.6	26.0	24.2	1.8	20.9	3.1	17.8	1.6	-0.6	2.2	51.5	44.7	6.8
12	0.103	74.7	25.3	22.5	16.4	6.1	58.1	43.5	14.6	0.2	-0.6	0.8	19.2	15.4	3.8
14	0.109	80.0	20.1	35.3	30.0	5.3	53.6	41.4	12.2	-1.3	-2.0	0.7	12.5	10.6	1.9
18	0.114	95.9	4.2	42.0	39.4	2.6	56.3	53.6	2.7	-2.0	-1.6	-0.4	3.8	4.5	-0.7
6	0.187	96.6	3.4	51.9	48.0	3.9	48.1	47.6	0.5	-2.1	-1.7	-0.4	2.1	2.7	-0.6
13	0.225	76.7	23.4	29.6	23.4	6.2	58.6	45.0	13.6	2.0	0.8	1.2	9.9	7.5	2.4
17	0.261	78.7	21.4	22.3	16.7	5.6	69.1	56.3	12.8	3.8	2.4	1.4	4.9	3.3	1.6

Source: Author's calculations

Remarks: $\nabla P_f = \nabla E_f + \nabla I_f$

$= \nabla P_{fc} + \nabla P_{fa} + \nabla P_{fh} + \nabla P_{fo}$

$= \nabla E_{fc} + \nabla E_{fa} + \nabla E_{fh} + \nabla E_{fo} + \nabla I_{fc} + \nabla I_{fa} + \nabla I_{fh} + \nabla I_{fo}$

$\nabla P_{fc} = \nabla E_{fc} + \nabla I_{fc}$, etc.

Table 6

Distribution of villages in states/Union Territories by Node/Cluster

Country/State/Union Territory	Node/Cluster										All
	6	9	10	12	13	14	15	16	17	18	
Jammu & Kashmir	9.2	16.1	17.7	22.1	1.7	5.4	3.4	6.3	0.7	17.4	100
Himachal Pradesh	4.0	15.3	0.9	16.0	2.8	2.7	35.2	22.1	0.2	0.8	100
Punjab	0.0	8.2	7.5	0.0	4.8	16.5	12.2	50.8	0.0	0.0	100
Chandigarh	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100
Uttarakhand	1.9	28.1	5.5	7.8	8.2	7.5	30.5	9.5	0.1	0.9	100
Haryana	0.0	9.3	40.0	0.0	1.7	16.2	1.5	31.4	0.0	0.0	100
Delhi	0.0	3.9	45.6	0.0	0.0	0.0	1.0	49.5	0.0	0.0	100
Rajasthan	15.3	19.9	15.8	11.1	2.1	6.8	0.6	1.7	4.1	22.7	100
Uttar Pradesh	0.3	25.4	49.5	5.9	1.3	7.2	2.2	5.9	0.1	2.4	100
Bihar	1.4	23.5	56.0	7.7	0.3	1.1	0.4	1.3	0.1	8.2	100
Sikkim	19.4	0.5	0.0	78.2	0.2	0.0	0.2	0.0	1.0	0.5	100
Arunachal Pradesh	85.8	2.5	0.6	5.7	0.1	0.1	0.1	0.1	0.8	4.2	100
Nagaland	94.7	0.1	0.0	4.0	0.0	0.0	0.1	0.0	0.3	0.9	100
Manipur	79.5	3.6	3.7	4.8	0.3	1.0	4.8	1.5	0.3	0.4	100
Mizoram	98.6	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	1.0	100
Tripura	49.1	0.3	2.9	39.1	0.0	0.0	1.0	7.1	0.0	0.5	100
Meghalaya	95.4	1.0	0.7	1.8	0.0	0.1	0.0	0.1	0.0	0.8	100
Assam	22.9	13.7	18.1	23.2	0.5	2.0	5.0	9.5	0.6	4.5	100
West Bengal	7.8	5.9	9.8	36.3	1.8	5.2	3.6	21.3	3.5	4.9	100
Jharkhand	36.1	18.0	10.0	13.5	0.4	1.4	0.4	1.0	1.2	18.0	100

Country/State/Union Territory	Node/Cluster										All
	6	9	10	12	13	14	15	16	17	18	
Odisha	30.1	4.7	1.1	25.3	1.2	2.7	7.7	14.3	3.8	9.1	100
Chhattisgarh	39.2	3.7	3.7	38.5	0.2	1.2	0.6	0.5	1.9	10.5	100
Madhya Pradesh	24.6	13.4	8.7	30.6	0.8	3.4	0.6	1.7	1.6	14.5	100
Gujarat	24.9	6.5	17.5	23.8	1.6	5.5	3.1	12.3	0.5	4.2	100
Daman and Dieu	15.8	0.0	15.8	63.2	0.0	0.0	0.0	5.3	0.0	0.0	100
Dadra & Nagar Haveli	93.8	0.0	0.0	6.2	0.0	0.0	0.0	0.0	0.0	0.0	100
Maharashtra	13.5	3.8	5.1	53.3	2.4	2.8	7.2	9.5	0.6	1.6	100
Andhra Pradesh	17.8	1.9	2.3	19.7	12.8	11.6	1.7	1.6	25.4	5.1	100
Karnataka	2.4	3.7	4.1	43.1	11.4	6.0	8.0	7.6	6.0	7.9	100
Goa	7.7	0.6	0.0	49.7	0.0	0.3	19.6	22.1	0.0	0.0	100
Lakshadweep	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100
Kerala	0.1	0.0	2.8	48.7	0.0	0.0	47.3	1.0	0.2	0.0	100
Tamil Nadu	2.0	1.7	1.9	19.2	18.2	14.1	25.0	16.3	1.3	0.4	100
Puducherry	0.0	0.0	3.3	0.0	0.0	4.4	85.6	6.7	0.0	0.0	100
Andaman & Nicobar Islands	19.4	17.2	1.6	16.9	2.2	0.6	6.9	35.3	0.0	0.0	100
India	15.6	13.0	18.1	21.2	2.9	5.1	5.4	8.5	2.7	7.4	100

Source: Author's calculations.

Figure 1
Distribution of villages by index of female participation, P_f, and occupation category

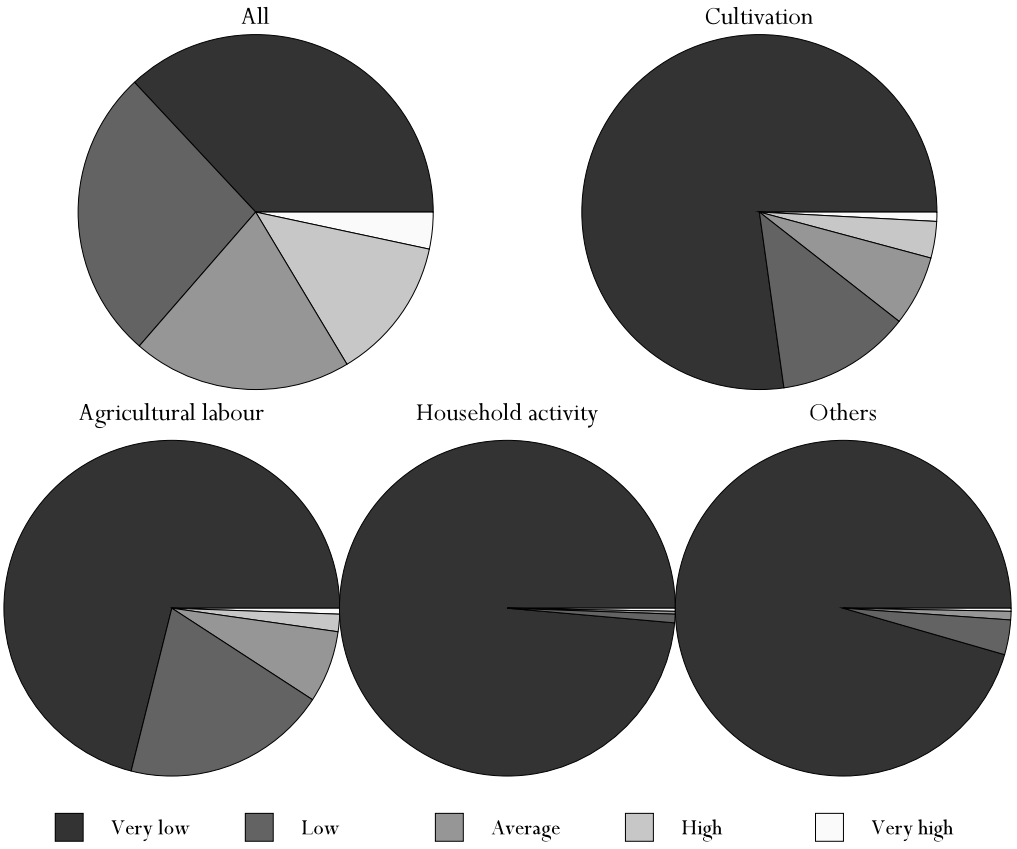

Figure 2
Classification of villages by their defining characteristics

Remarks:

Social class structure is measured in terms of the proportion of Scheduled Tribes females to total females in the village.

Fertility is measured in terms of the ratio of children aged 0-6 years to females aged 7 years and above.

Gender composition is measured in terms of the ratio of males aged 7 years and above to females aged 7 years and above.

Female literacy is measured in terms of the proportion of females aged 7 years and above who can read and write with understanding.