# Chapter VI

## Covariates of Rural Female Work Participation - A Logistic Regression Exercise

### 6.1. INTRODUCTION

Women's labour market decisions along with the type of work to be done cannot be made by women alone (Atal, 2017) since such decisions depend strongly on the decisions taken at the household level (ILO, 2012, p.35) and are determined by a myriad of factors. Within the household, members of the family divide paid work, unpaid housework (including cooking, cleaning, and washing) and care work (care of small children and the elderly) among themselves to fulfill the household needs. This division of labour among the members of the household is determined by several factors which include individual notions regarding gender roles, child rearing practices, perceptions regarding value of women's contributions etc. These are in turn influenced by household members' relative bargaining power and are affected by "potential income, human capital, economic dependency, potential status in employment along with specific household needs and interests" (*ibid*).

Despite contributing significantly to the functioning of an economy, female labour force is one of the most underutilized and neglected human resource, particularly in the Third World nations which makes them disadvantaged socially and economically and has important implications for economic welfare and growth (Psacharopoulos and Tzannatos, 1989). The Human Development Report, 2015 notes that with regard to work-both unpaid care work and paid work, there are significant imbalances between men and women with women performing three times more than men when unpaid work is considered, while men's share is nearly twice that of women in case of paid work mostly outside the home (UNDP, 2015, pp. 11-12). The higher involvement of women in unpaid work is not a question of their choice, but is determined by societal norms regarding the gender division of labor which considers men to be the breadwinner and women to be the "home-maker, mother and dependent" (ActionAid 2017, p. 15). Since it does not have any exchange value, unpaid work escapes statistical visibility as labour force statistics of most countries measure only paid work. Further, many women do not consider themselves as being employed, if they are not earning wages in cash (Sundar, 1981).

It has been argued that in contrast to women who are involved only in domestic work, women who earn wage income enjoy a higher status within and outside the household and have increased bargaining power which may be attributed to the lower valuation of unpaid

household work. Women's participation in paid work may nonetheless not always be empowering. As pointed out by Chakraborty and Chakraborty (2009) increased participation by women during economic distress may cause young girls to drop out of school to help in domestic chores and sibling care which restricts their schooling, and in turn leads to a widening of the gender gap in education and in labour market opportunities. This implies that women, due to lower levels of education may be employed in jobs that have low productivity and are casual in nature. Gender segregation within the household is therefore translated into gender segregation in the labour market.

Labour force participation of women is known to be a function of multifarious factors. More than economic, non-economic factors have been found to be significant in explaining the labour market behaviour of women. Besides economic variables such as education, experience, wages, and incomes, scholars argue that female labor supply may be determined by non economic variables such as marital status and fertility, urbanization, ownership of land and size of farm, status of household head, and structure of employment (Psacharopoulos and Tzannatos, 1989). Being determined by such varied factors there is considerable heterogeneity in the female labour force participation rates between different regions and nations. As appropriately mentioned by Standing (1981) since the level, patterns and trends in female labor force participation vary widely between and within countries it would be misleading to make any generalizations (Standing, 1981 cited in *ibid*). Female labour force participation should therefore be analysed with reference to the socio-economic and demographic characteristics of the area under study.

With this objective in view, the present chapter looks into the contribution of rural women to the hill economy in terms of their work participation in paid work and unpaid work in family farms. A theoretical framework is initially formulated to understand the work contributions of the rural women to the food and economic security and the welfare of the households. Thereafter, using binary logistic regression analysis the study isolates the determinants of women's work participation in the study area.

## 6.2. WOMEN'S WORK AND THEIR CONTRIBUTIONS TO HOUSEHOLD FOOD AND ECONOMIC SECURITY- A THEORETICAL FRAMEWORK

Following Sidh and Basu (2011) a theoretical framework may be formed for understanding women's work in the study area and their contribution to fulfilling the family's requirements of food, economic and non-economic goods for ensuring the food and economic security of the household as shown in Figure 6.1. The households comprise adult male and female

members along with children or older family members. Women of the region perform multiple tasks for ensuring food and economic security of the household. They are responsible for family subsistence and in many instances such as female headed households they are the primary or sole economic providers (Agarwal, 1989). The activities performed by women in the study area may be broadly classified into four categories- crop production; livestock rearing; unpaid household activities which includes collection of fuel wood, fodder for animals, water and non-timber forest products; and participation in paid activities.

Agricultural production is a function of the amount of land owned by the households, along with farm animals, farm implements and the inputs of male and female labour-own or hired. Women in the region play an important role in crop production by participating in most activities either as principal or subsidiary status workers. Agriculture is the region is primarily for subsistence although in certain households it is carried out for commercial purposes. Most of the families grow their own vegetables and also paddy in some cases, most of which is used for self consumption. Households also derive income from sale of agricultural products which include cash crops such as black cardamom, broom grass, red round chillies, and some seasonal vegetables like squash, green leafy vegetables etc. Households use inputs of family labour-male, female and child for maintenance of farm animals like cows, goats, pigs, poultry etc. which is also used to fulfill the family's requirements for milk and dairy products, eggs, meat etc. and also supplement family income when these products are sold in the market. Livestock, particularly bovines supply manure and draught power in crop production activities. The labour inputs of women in crop production and livestock rearing thus contribute substantially to household food security directly by supplying food articles, and indirectly through the sale of these products in the market, thus ensuring food and economic security of the households.

Typical of rural economy, women in the region are burdened with the responsibility of running the household which includes activities that lie outside the purview of conventional economic activities. Household tasks are plenty and include preparing food, washing and cleaning, maintaining the households, purchasing household articles and looking after the children and the elderly. Women perform these tasks with little or no help from the men. Collection of fodder for animals, fuel wood, water and non timber forest products require the labour inputs of men, women and sometimes children. Household work, the disproportionate burden of which is borne by the women is indispensible for the proper functioning of the households and therefore contributes significantly to the welfare and food

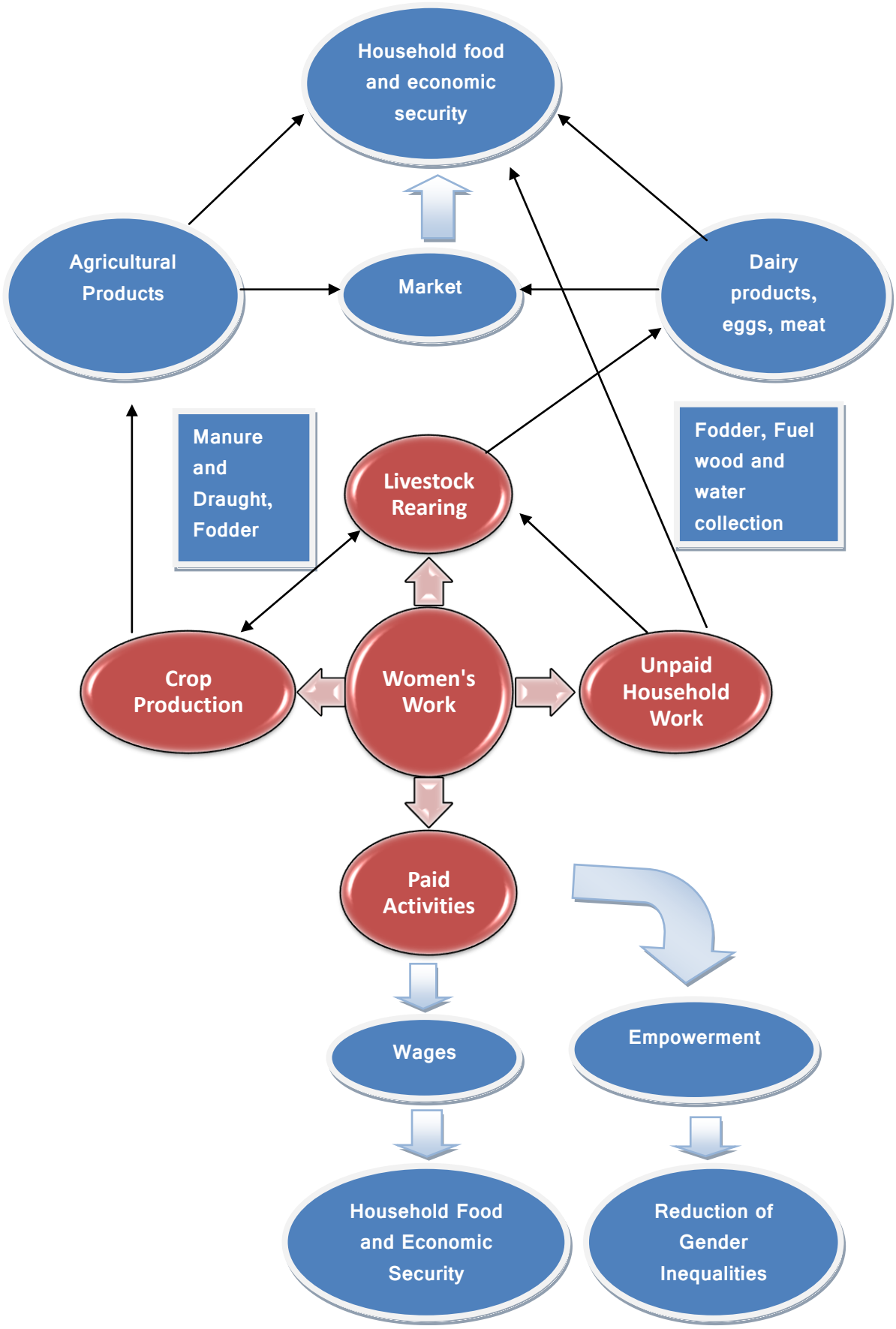


**Figure 6.1: Women's Work and their Contribution to Food and Economic Security**

security of the households. Besides, women are also involved in building up of family networks and inter-household relationships in the village, which often prove to be of critical importance for survival during periods of food shortages (Agarwal, 1988a cited in Agarwal, 1989).

In comparison to men, women's ability to participate in economic activities is less and constrained by the burden of household work and child care which men do not face. Being time intensive, household and child care responsibilities restrict women's mobility outside the home and limit the time spent by women in other income generating activities. Participation in paid activities which generally includes non-farm activities or farm activities as paid agricultural labour on somebody else's farms directly contributes to household food and economic security by enhancing the household income. In addition it makes women socially and economically empowered ultimately leading to reduction in gender inequalities. On the contrary, participation in paid activities may increase women's work burdens since they now have to work both within and outside the households leading to a dual burden (Choudhary and Parthasarathy, 2007) since the participation of men in so-called feminine indoor tasks is not evident (Bittman and Matheson, 1996 cited in *ibid*) even when women participate in activities outside the household.

The realization of women's responsibility in achieving food and economic security for the family is however limited by their unequal access to resources-an inequality arising not by virtue of their class but by virtue of their gender. These inequalities may be in the form of differences in the distribution of basic necessities within the family, disadvantaged position of women in the labour market, women's limited access to means of production such as land and production technology along with deterioration and privatisation of common property resources which are critical for the sustenance of the poor, particularly women (Agarwal, 1989).

## 6.3. DETERMINANTS OF RURAL FEMALE WORK PARTICIPATION-A LOGISTIC REGRESSION EXERCISE

Although sufficient research exists on women's work participation and its determinants in different regions of the country, literature on women's work participation rates in the hill and mountain regions of the country which is relatively high is somewhat limited. As per the Census 2011 data, the hill district of Darjeeling (including Kalimpong sub-division) in West Bengal ranked third among all the districts for rural female WPR with a value of 26 per cent, which was also higher than the state average of 19.4 per cent. In the hill regions agricultural

activities are often conducted for subsistence and are distinctly different from those practiced in the plains, due to constraints imposed by altitude and topography. Women's participation in economic activities in such settings is affected by several factors that have received little attention from scholars. Keeping this in mind, the present section attempts to identify the determinants of work participation of rural women in the hill region in the district of Darjeeling (including Kalimpong sub-division).

The Employment and Unemployment Survey of the National Sample Survey Organisation (NSSO), 68th Round, defines economic activity as any activity which leads to the production of goods and services and adds to the national product. It includes "(i) production of all goods and services for the market (i.e. for pay or profit) including government services, (ii) production of primary commodities for own consumption, and (iii) own account production of fixed assets" (GoI, 2013). In the present study, the work participation of women is based only on *usual principal activity status*, i.e. a woman is said to be in the workforce according to the *usual status (ps)* if she participates in any kind of activity for the major part of the 365 days preceding the survey. Women who are primarily engaged in household work and participate in economic activities according to the *usual subsidiary activity status (ss)* are considered to be outside the workforce. Unemployed women are not included in the workforce. The workforce therefore includes women in paid employment and those who work as unpaid family labour on family farms (Rai and Mukherjee, 2018). Women who participate in economic activities according to *subsidiary activity status* have not been considered in the present analysis since women (and men) in rural areas are engaged in multiple occupations on a subsidiary basis to augment family income. Even women who are out of the workforce as per *principal activity status* may be engaged in some form of economic activity such as family labour on family farms or involved in animal husbandry if *subsidiary activity status* is taken into account. Inclusion of workers according to *subsidiary activity status* would lead to very less/no participants in the category of non-workers, and have thus been excluded from the present analysis.

Following Ackah, Ahiadeke and Fenny (2009), two alternative models have been estimated in the present analysis on the basis of female workforce participation. Model I takes into consideration both paid and unpaid employment of women as participation in the workforce. It includes women in paid/wage work or self-employment, as well as those engaged in family farms as unpaid family labour. In order to arrive at the determinants of women's paid work, Model II considers a more restrictive definition of participation that includes only paid market work, but covers both wage work and self-employment. Paid

market work may be distinguished from unpaid work as "work that is remunerated in cash or kind in the shape of wages, salaries and profit" whereas unpaid work is work "performed without any direct remuneration" (Actionaid 2017, p. 13). The data were analysed using Statistical Package for the Social Sciences (SPSS) version 23 and Econometric Views (EViews) version 10 (Rai and Mukherjee, 2018).

### 6.3.1. Model Specification

The decision to participate or not to participate in the work force is a binary one and can take only two values i.e. either a "yes" or a "no". Given the dichotomous nature of the response variable binary logistic regression has been used in the present analysis. In the binary logistic regression model, the female work force participation which is the dependent variable (Y) can take only two values, $Y_i = 1$ if the respondent is in the workforce and $Y_i = 0$ if the respondent is not in the workforce. The dependent variable is determined by the predictor variables $X_i$s also known as covariates which may be numerical or categorical in nature. For a categorical variable, dummy variable is used to compare the different categories. For each categorical variable a baseline or reference category is chosen and all other categories are compared to the baseline category. For a categorical variable with k categories, k-1 dummy variables are to be introduced (Gujarati 2004, p. 302).

If P$i$ represents the probability for the $i^{th}$ female respondent to be in the workforce, the model may be written as

Probability [A female is the workforce] = P$i$ $(Y = 1) = \frac{1}{1+e^{-Z}} = \frac{e^{Z}}{1+e^{z}}$ ----------(1)

where Z is a linear function of the explanatory variables. If $X_1$, $X_{2,\ldots\ldots}X_k$ represent the various explanatory variables, then "Z" equation would be

$Z_i = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_k X_k$

$X_i$ = $i^{th}$ Explanatory variables (i = 1,2, ….., k) and

$\beta_i$ = Parameters of the model (i = 0, 1, 2, ….. , k)

This is known as the cumulative logistic distribution function. From the above equation it can be seen that the P$i$ is nonlinear not only in $X_i s$ but also in the β's. Hence the Ordinary Least Squares (OLS) procedure cannot be used to estimate the parameters of the equation. However it can be linearized in the following manner. If P$i$ is the probability of a female being in the workforce, then (1- P$i$) is the probability of the female not being in the workforce which may be written as

Probability [A female is not in the workforce] = 1- P$i$ (Y = 1)

$$=1-\frac{1}{1+e^{-Z}}$$

$$=\frac{1}{1+e^{Z}} \text{ ----------(2)}$$

Now we can write $\frac{P_i}{1-P_i}=\frac{1+e^{Z}}{1+e^{-Z}}=e^{Z}$ ----------(3)

Here $\frac{P_i}{1-P_i}$ are the odds in favour of a respondent participating in the workforce which is simply the ratio between the probabilities of an event occurring to the probability of not occurring. Taking the natural logarithm of the odds we get

$$L_i(\text{logit})=\ln\left(\frac{P_i}{1-P_i}\right)=Z_i \text{ ----------(4)}$$

$$L_i=\ln\left(\frac{P_i}{1-P_i}\right)=\beta_o+\beta_1 X_1+\beta_2 X_2+\ldots..+\beta_k X_k \text{ ----------(5)}$$

It can be seen that $L_i$ the log of the odds is linear not only in $X_i s$ but also in the β's and is called the logit, hence the name logit model (Gujarati 2004, pp. 595-596). It can also be seen that as the value of $P_i$ changes from 0 to 1, the odds change from 0 to ∞. When $P_i$ = 0.5, the odds are 1. On the odds scale the values from 0 to 1 correspond to values of $P_i$ from 0 to 0.5. On the other hand, values of $P_i$ from 0.5 to 1.0 result in odds of 1 to ∞. This asymmetry can be corrected by taking the natural logarithm of the odds. When $P_i$ = 0, ln (odds) = -∞; when $P_i$ = 0.5, ln (odds) = 0.0; and when $P_i$ = 1.0, ln (odds) =+∞ (Afifi, May and Clark, 2012, p. 272). This implies that although the probabilities lie between 0 and 1, the logit is not bounded and can lie between -∞ and +∞, as such it can have an unlimited range of values. The link function (the function of the dependent variable that yields a linear function of the independent variables) in logistic regression model is therefore the logit transformation (Hosmer and Lemeshow, 2000 pp. 48).

Many of the assumptions of the linear regression models do not apply to logistic regression analysis such as linearity of relationship between the dependent and independent variables, normality of the error distribution, homoscedasticity of the errors, and measurement level of the independent variables. Since logistic regression applies a non-linear log transformation of the linear regression, non-linear relationships between the dependent and independent variables can also be considered in logistic regression. It can also handle continuous and discrete data as independent variables which is a principal advantage of logistic regression model (Park, 2013).

Though the log odds of the dependent variable can be expressed as a linear combination of the predictors, the estimation of the logistic regression through the method of least squares as in linear regression yields estimators with a number of undesirable statistical properties. The most preferred method of estimation for a logistic regression model is maximum likelihood estimation. Estimation under this method requires construction of a function called the likelihood function which expresses the probability of the observed data as a function of the unknown parameters i.e. the βs. The maximum likelihood estimators of these parameters are chosen to be those values that maximise this function i.e. the estimated parameters are those which maximise the probability of obtaining the observed data (Hosmer and Lemeshow, 2000, p. 8).

In logistic regression the results are obtained in the form of the log of odds which is the slope coefficients i.e. the $\beta_i$s which measures the change in log odds or logit of participating in the work force for a unit change in the predictor variable holding the other predictor variables constant. In logistic regression, proper interpretation of the coefficient depends on being able to meaningfully explain the difference between two logits (*ibid*, p.48). For a more meaningful interpretation of the results of the logistic regression, the log odds can be converted into the odds ratio by exponentiating the logarithmic values. The odds ratio is a ratio between two odds i.e. the ratio of odds of occurrence of an event between two situations. If we consider two values of the categorical independent variable, X i.e. X=0 and X=1, then the odds ratio for Y=1 may be written as

Odds Ratio = $\frac{P_1/_{1-P_1}}{P_0/_{1-P_0}}$ where $P_1$ is the probability of Y=1 for X=1 and $P_0$ is the probability of occurrence of Y=1 for X=0. The numerator therefore represents the odds in favour of the event Y for X=1 and the denominator represents the odds in favour of Y for X=0.

Taking one independent variable X, it can be seen from equation (3) that

$\frac{P_i}{1-P_i} = \frac{1+e^Z}{1+e^{-Z}} = e^Z$ where $Z = \beta_o + \beta_1 X$

So the Odds Ratio for X=1 and X=0 can be written as

Odds Ratio= $\frac{P_1/_{1-P_1}}{P_0/_{1-P_0}}$

$= \frac{e^{\beta 0 + \beta 1}}{e^{\beta 0}}$

$= e^{\beta_1}$ ---------- (6)

It can be seen from equation (6) that the exponentiated coefficient shows the ratio of two odds. This simple relationship between the odds ratio and the coefficient is the primary reason why logistic regression is considered to be such a powerful research tool in analytical studies (*ibid*, pp.48-50). Exp (β) is therefore the incremental odds ratio and corresponds to a one unit increase in the variable X, assuming that values of all the other X variables does not change. The incremental odds ratio corresponding to the change of *k* units in X is exp (*k*β) (Afifi et. al., 2012, p. 276). A unit increase in the predictor increases or decreases the odds in favour of the event i.e. female work force participation accordingly as the odds ratio is greater than or less than 1 (Rai and Mukherjee, 2018).

### 6.3.2. Variables Used in the Model

A review of the literature shows that various socio-economic variables influence the work participation behaviour of women. Table 6.1 lists the explanatory variables chosen for analysis along with the expected sign. Age squared has been used to capture the non-linear effect of age.

**Table 6.1: Variables Used in the Model (Logistic Regression)**

| Variables | Notation | Description | Expected Signs |
|---|---|---|---|
| **Dependent Variable** | | | |
| Model I: Female work participation | FLFP | Dummy Variable =1 if participating in the work force<br>=0 otherwise | |
| Model II: Female work participation in paid work | FLFP_PAID | Dummy Variable =1 if participating in the paid workforce<br>=0 otherwise | |
| **Independent Variables** | | | |
| Age | AGE | Number of years completed | Positive |
| Age squared | AGE_SQU | Square of the number of years completed | Negative |
| Education | EDUCATION | Number of years of schooling | Positive |
| Family Structure | FAM_STR | Dummy Variable =1 Joint<br>= 0 Unitary | Positive/Negative |
| Number of children below the age of six years | CHILD_06 | Dummy Variable =1 if child below the age of six years is present<br>= 0 otherwise | Negative |
| Woman's marital status | MARITAL_STATUS | Dummy variable=1 if currently married<br>=0 otherwise | Negative |
| Primary occupation of household head | OCCUPATION_HEAD | Dummy Variable=1 if primary occupation is agriculture<br>=0 otherwise | Positive |
| Presence of male migrant member | MIGRANT | Dummy Variable=1 if migrant male member is present<br>=0 otherwise | Positive/Negative |
| Monthly per capita consumption expenditure | MPCE | Monthly per capita consumption expenditure of household in Rs 1,000 | Negative |
| Land holding | LAND | Ownership holding of the household in acres | Positive/Negative |

It is however, very likely that the age and age-squared terms are highly correlated. This is however not something to be concerned about because the *p*-value of the squared term is not affected by multicollinearity. The high correlation can be reduced by centering the variables (i.e. by subtracting the means) before squaring which will have no effect on *p*-value of the squared term and results for other variables including the R squared, but not the lower order terms which means that multicollinearity is not a problem (Allison, 2012). In the present analysis age has been centered to reduce multicollinearity. The monthly per capita consumption expenditure of the household has been used as a proxy for household income to avoid the problem of endogeneity in the model. The total expenditure incurred by a household on domestic consumption during the reference period is known as the household's consumption expenditure (GoI, 2014, p. 8). In the present model, household expenditure on food items was estimated per month. For self-produced items such as vegetables and milk, the quantity consumed per month was recorded and the value estimated according to the prevailing market price. Expenditure on other items such as clothing, education, and agricultural inputs was estimated using a one-year recall period. Recall errors are an inherent limitation of the present study. Household monthly per capita consumption expenditure (MPCE) was measured as a continuous variable in thousand rupees (Rai and Mukherjee, 2018).

### 6.3.3. Selection of Variables

The criteria for inclusion of a variable in a model differs between problems and scientific disciplines. Traditionally, statistical model building involves looking for the most parsimonious model that still explains the data, the rationale for minimising the number of variables in the model being that the resultant model is more likely to be numerically stable, and can be more easily generalised. Some methodologies suggest inclusion of all intuitively relevant variables in the model not considering their statistical significance in order to control for confounding factors. This approach can however, lead to numerically unstable estimates characterised by unrealistically large estimated coefficients and/or large standard errors (Hosmer and Lemeshow 2000, p. 92). For a purposeful selection of variables, Hosmer and Lemeshow suggest carrying out a univariable analysis of each variable (*ibid.*). After completion of the univariable analysis, the significance of the univariate analysis can be examined by comparing the *p* values for the different explanatory variables to some arbitrary level, which is 0.25. Any variable whose univariable test has a *p*-value < 0.25 can be included

in the multivariable model (*ibid*, p. 95). The univariate test for both models has been presented in Tables 6.2 and 6.3 (Rai and Mukherjee, 2018).

**Table 6.2: Univariate Analysis for Model I (Dependent Variable- FLFP)**

| FLFP | | B | S.E. | Wald | df | Sig. | Exp(B |
|---|---|---|---|---|---|---|---|
| Step 1a | AGE | 0.106 | 0.018 | 33.07 | 1.000 | 0.000 | 1.112 |
| Step 1a | AGE_SQU | -0.004 | 0.001 | 16.70 | 1.000 | 0.000 | 0.996 |
| Step 1a | EDUCATION | -0.080 | 0.032 | 6.183 | 1.000 | 0.013 | 0.923 |
| Step 1a | FAM_STR(1) | -0.845 | 0.306 | 7.633 | 1.000 | 0.006 | 0.430 |
| Step 1a | CHILD_06(1) | -1.221 | 0.369 | 10.95 | 1.000 | 0.001 | 0.295 |
| Step 1a | MARITAL_STATUS(1) | 1.062 | 0.321 | 10.96 | 1.000 | 0.001 | 2.894 |
| Step 1a | OCCUPATION_HEAD(1 | 0.965 | 0.299 | 10.41 | 1.000 | 0.001 | 2.624 |
| Step 1a | MIGRANT(1) | -0.660 | 0.334 | 3.920 | 1.000 | 0.048 | 0.517 |
| Step 1a | MPCE | 0.034 | 0.081 | 0.175 | 1.000 | 0.676 | 1.034 |
| Step 1a | LAND | -0.025 | 0.080 | 0.099 | 1.000 | 0.754 | 0.975 |

a. Variable(s) entered on step 1: AGE/ AGE_SQU/EDUCATION/ FAM_STR/ CHILD_06/ MARITAL_STATUS/ OCCUPATION_HEAD/ MIGRANT/MPCE/LAND

On the basis of the criteria for selection of variables as proposed by Hosmer and Lemeshow, MPCE and LAND have been excluded as explanatory variables for Model I. For Model II, AGE and MPCE have *p* values < 0.25. However, as AGE_SQU is statistically significant, both AGE and AGE_SQU have been included as explanatory variables. MPCE has been excluded due to a *p* value < 0.25 (Rai and Mukherjee, 2018).

**Table 6.3: Univariate Analysis for Model II (Dependent Variable- FLFP_PAID)**

| FLFP_PAID | | B | S.E. | Wald | df | Sig. | Exp(B) |
|---|---|---|---|---|---|---|---|
| Step 1a | AGE | -0.008 | 0.017 | 0.238 | 1.000 | 0.625 | 0.992 |
| Step 1a | AGE_SQU | -0.004 | 0.002 | 4.196 | 1.000 | 0.041 | 0.996 |
| Step 1a | EDUCATION | 0.084 | 0.038 | 5.056 | 1.000 | 0.025 | 1.088 |
| Step 1a | FAM_STR(1) | -1.186 | 0.403 | 8.686 | 1.000 | 0.003 | 0.305 |
| Step 1a | CHILD_06(1) | -1.974 | 1.032 | 3.662 | 1.000 | 0.056 | 0.139 |
| Step 1a | MARITAL_STATUS(1) | -0.477 | 0.403 | 1.401 | 1.000 | 0.237 | 0.621 |
| Step 1a | OCCUPATION_HEAD(1) | -0.507 | 0.374 | 1.839 | 1.000 | 0.175 | 0.602 |
| Step 1a | MIGRANT(1) | -1.176 | 0.627 | 3.520 | 1.000 | 0.061 | 0.308 |
| Step 1a | MPCE | 0.014 | 0.100 | 0.020 | 1.000 | 0.888 | 1.014 |
| Step 1a | LAND | -0.649 | 0.246 | 6.962 | 1.000 | 0.008 | 0.523 |

a. Variable(s) entered on step 1: AGE/ AGE_SQU/EDUCATION/ FAM_STR/ CHILD_06/ MARITAL_STATUS/ OCCUPATION_HEAD/ MIGRANT/MPCE/LAND

## 6.4. REPORTING THE LOGISTIC REGRESSION RESULTS

### 6.4.1. Frequency Table of Binary Responses

The frequencies of the binary response variable i.e. female work force participation for Model I and Model II have been represented graphically (Figure 6.2) and in tabular form (Table 6.4) below.

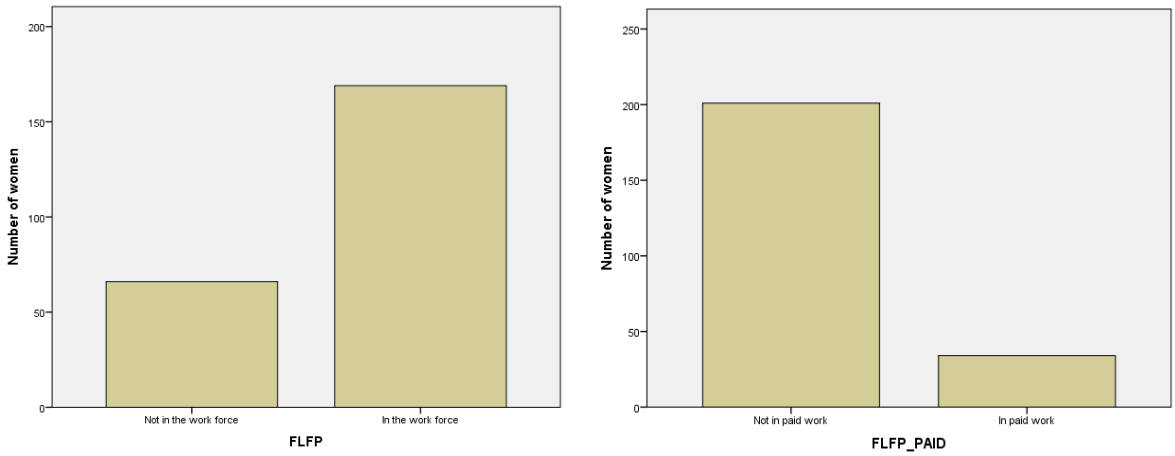


**Figure 6.2: Frequency of Binary Response Variable (Model I: Dependent Variable-FLFP and Model II: Dependent Variable-FLFP_ PAID)**

The number of households surveyed in each of the three villages was 50 leading to a total of 150 households surveyed. Among the surveyed houeholds only women between the ages of 15 and 65 were considered for the analysis, irrespective of whether or not they participated in an economic activity. Females attending educational institutions were excluded from the analysis. This gave a total sample size of 235 (Samalbong village-68; Git Dubling Khasmahal-74; Sitong Khasmahal-93).

**Table 6.4: Frequency Table for the Binary Response Variable (Model I Dependent Variable-FLFP and Model II Dependent Variable-FLFP_ PAID)**

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Model I | | | | | |
| Valid | FLFP=0 | 66 | 28.1 | 28.1 | 28.1 |
| | FLFP=1 | 169 | 71.9 | 71.9 | 100 |
| | Total | 235 | 100 | 100 | |
| Model II | | | | | |
| Valid | FLFP_PAID=0 | 201 | 85.5 | 85.5 | 85.5 |
| | FLFP_PAID=1 | 34 | 14.5 | 14.5 | 100 |
| | Total | 235 | 100 | 100 | |

Among 235 working aged women, 169 (71.9 percent) are participating in the work force as opposed to 66 (28.1 percent) who are participating only in domestic activities according to Model I. According to Model II which takes into account women's paid activities, 34 (14.5 percent) are participating in the work force compared to 201 (85.5 percent) who are participating either as unpaid family labour on family farms or are engaged in domestic activities only (Rai and Mukherjee, 2018).

### 6.4.2. Summary Statistics

Table 6.5 shows the mean and standard deviation of the explanatory variables used in the study for the two models. The average age of women in the work force in Model I (40.97) is higher than that in Model II (37.29). However, the level of education as measured by years of schooling for women who are in the work force is higher for Model II (8.71) than for Model I (6.11). The average size of landholding of women in the work force in Model II is lower than that in Model I, whereas the average monthly per capita consumption expenditure of women who are in the workforce is almost equal in the two models. The average monthly per capita consumption expenditure for women who are in the work force for Model I is Rs 2,867.81 and Rs 2,877.35 for Model II; and the average size of landholding is 1.97 acres for Model I and 1.26 acres for Model II. For the categorical explanatory variables, the means indicate the proportion of cases with value of the explanatory variable = 1 in each category of the dependent variable. For example, a mean of 0.48 for FAMILY_STR for FLFP = 1 implies that 48 per cent of women in the work force in Model I are in joint families (Rai and Mukherjee, 2018).

**Table 6.5: Summary Statistics of Sample**

| Explanatory Variables | Model I | | | | Model II | | | |
|---|---|---|---|---|---|---|---|---|
| | FLFP=1 N=169 | | FLFP=0 N= 66 | | FLFP_PAID=1 N=34 | | FLFP_PAID=0 N=201 | |
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| AGE | 40.97 | 9.38 | 30.94 | 11.98 | 37.29 | 8.59 | 38.3 | 11.5 |
| EDUCATION | 6.11 | 5.19 | 8.08 | 4.39 | 8.71 | 7.91 | 6.32 | 4.31 |
| FAMILY_STR | 0.48 | 0.50 | 0.68 | 0.47 | 0.29 | 0.46 | 0.58 | 0.50 |
| CHILD_06 | 0.11 | 0.31 | 0.29 | 0.46 | 0.03 | 0.17 | 0.18 | 0.38 |
| MARITAL_STATUS | 0.82 | 0.39 | 0.61 | 0.49 | 0.68 | 0.48 | 0.77 | 0.42 |
| OCCUPATION_HEAD | 0.62 | 0.49 | 0.38 | 0.49 | 0.44 | 0.50 | 0.57 | 0.50 |
| MIGRANT | 0.18 | 0.39 | 0.30 | 0.46 | 0.09 | 0.29 | 0.24 | 0.43 |
| MPCE | 2.87 | 1.72 | 2.76 | 2.11 | 2.88 | 1.56 | 2.83 | 1.88 |
| LAND | 1.97 | 1.73 | 9.34 | 1.92 | 1.26 | 0.87 | 2.12 | 1.87 |
| Sample Size | 235 | | | | 235 | | | |

## 6.5. EVALUATION OF THE LOGISTIC REGRESSION MODEL

### 6.5.1. Likelihood Ratio Test

The overall fit of the model can be assessed by comparing the fit of the intercept-only model or the null model with the model containing the independent variables. A logistic regression model with a given number of independent variables (the given model) is said to provide a better fit to the data if it reveals an improvement over the model with no independent variables (the null model) (Park, 2013). The overall fit of the model can be examined via a likelihood ratio test. It measures how the explanatory variables improve the fit of the given model compared to the null model.

The likelihood ratio test is based on –2log likelihood ratio in which a comparison is made between the deviance with just the intercept (-2 log likelihood of the null model) and the deviance when the *k* independent variables have been included (-2 log likelihood of the given model) (*ibid*). The likelihood of obtaining the observation if the independent variables had no effect on the result is the likelihood of the null model, whereas likelihood of the given model is the likelihood of obtaining the observations with all independent variables incorporated in the model (*ibid*). The Likelihood ratio statistic (LR) obtained from the –2log likelihood ratio yields a goodness of fit index $\chi 2$ statistic with *k* degrees of freedom (Bewick, Cheek, & Ball, 2005 cited in *ibid*). Significance at the .05 level or lower implies that the given model with the predictors is significantly different from the null model i.e. the constant only model which means that at least one of the independent variables contributes to the outcome.

**Table 6.6: Omnibus Tests of Model Coefficients-Model I and Model II**

| **Model I Dependent Variable=FLFP** | | | | | **Model II Dependent Variable=FLFP_PAID** | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **Chi-** | **df** | **Sig.** | | | **Chi-square** | **df** | **Sig.** |
| Step 1 | Step | 94.443* | 8 | 0.000 | Step 1 | Step | 46.268# | 9 | 0.000 |
| | Block | 94.443 | 8 | 0.000 | | Block | 46.268 | 9 | 0.000 |
| | Model | 94.443 | 8 | 0.000 | | Model | 46.268 | 9 | 0.000 |

**Initial -2 Log Likelihood=279.065, Model -2 Log Likelihood=184.622, hence LR statistic=279.065-184.622=94.443 #Initial -2 Log Likelihood=194.284, Model -2 Log Likelihood=148.016, hence LR statistic=194.284-148.016= 46.268*

The statistical software used for the present analysis i.e. SPSS version 23 provides the table for the Iteration History which shows the value of the initial -2 Log Likelihood and the - 2 Log Likelihood under the given model with predictors. The difference between the initial -2 Log Likelihood and the model -2 Log Likelihood gives the LR statistic which is referred to as

the Omnibus Tests of Model Coefficients and is shown in the Table 6.6. For both the models the model chi square turns out to be significant indicating that at least one of the predictors is significantly related to the outcome variable. In the present analysis since all the variables have been entered at the same time using the block entry of variables there is only one model to be compared with the null model, there is no difference in the results in step, block or model chi-square values (Rai and Mukherjee, 2018).

### 6.5.2. Hosmer and Lemeshow Goodness of Fit

Hosmer and Lemeshow proposed a goodness-of-fit test, now universally referred to as the Hosmer-Lemeshow test. The Hosmer-Lemeshow (H-L) statistic measures the difference between the observed and the predicted values of the dependent variable. Hosmer and Lemeshow proposed grouping of the observations into deciles on the basis of estimated probabilities (Hosmer and Lemeshow, 2000, p.147). This is shown by the Contingency Table for the Hosmer and Lemeshow Test shown in Appendix B. The H-L test statistic is then computed which asymptotically follows a $\chi 2$ distribution with 8 (number of groups -2) degrees of freedom (Park, 2013). The probability value ($p$) is also computed to test the goodness of fit of the model. A non-significant chi-square value for the H-L statistic ($p$>0.05) indicates a good fit model as it implies that the model prediction is not significantly different from the actual values. In the present analysis, as can be seen from Table 6.7 the H-L statistic for both the models yields a desirable outcome which is non significant ($p$=0.636 for Model I and $p$=0.557 for Model II) indicating that the predicted model does not significantly differ from the observed model (Rai and Mukherjee, 2018).

**Table 6.7: Hosmer and Lemeshow Goodness of Fit Test**

| **Model I Dependent Variable-FLFP** | | | | **Model II Dependent Variable-FLFP_PAID** | | | |
|---|---|---|---|---|---|---|---|
| **Step** | **Chi-square** | **df** | **Sig.** | **Step** | **Chi-square** | **df** | **Sig.** |
| 1 | 6.101 | 8 | 0.636 | 1 | 6.812 | 8 | 0.557 |

### 6.5.3. $R^2$ Equivalents for logistic regression/ Pseudo R-square

As compared to the $R^2$ in Ordinary Least Squares (OLS) regression which measures the overall fit of the linear regression model by explaining the proportion of variation in the dependent variable due to the independent variables, no such equivalent measure is present in logistic regression. However, a number of measures sometimes referred to as Pseudo R-square may be used as an equivalence of the $R^2$ in OLS although no consensus exists as to which one is the best. The two most commonly used measures of Pseudo R-square as reported by SPSS are Cox & Snell R Square and Nagelkerke R Square. The problem with

Cox and Snell's R square as suggested by Cox and Snell (1989) is that it does not reach the maximum value of 1 which makes it difficult to interpret. The Nagelkerke R-square is an improvement over the Cox and Snell's R square and a more reliable measure. Since its value ranges from 0 to 1 Nagelkerke's R square will normally be higher than the Cox and Snell measure. There is another measure that corresponds to Pseudo $R^2$ as reported by EViews which is the McFadden R squared whose value ranges from 0 to 1.

The Pseudo R square values for both the models have been presented in Table 6.8. It has been mentioned that there exists a direct empirical relationship between the $R^2$of a linear regression model and the pseudo $R^2$ of a choice model (Domencich and Mc Fadden, 1975) with pseudo $R^2$ values between the range 0.3 and 0.4 being translated as an $R^2$ of between 0.6 and 0.8 for the linear model equivalent (cited in Hensher, Rose and Green, 2005, pp. 338-339). Since the pseudo $R^2$ as shown by Nagelkerke R square is 0.476 and 0.318 for the two models the model fit is quite satisfactory for both the models (Rai and Mukherjee, 2018). Afifi et.al (2012) however notes that many statisticians consider pseudo $R^2$ to be of limited value since it does not represent the proportion of variation explained by the covariates and should not be interpreted as $R^2$ in multiple regression (Afifi et.al., 2012, p. 295).

**Table 6.8: Values of Pseudo $R^2$**

| **Pseudo $R^2$** | **Model I (Dependent variable – FLFP)** | **Model II (Dependent variable – FLFP_PAID)** |
|---|---|---|
| McFadden R-square | 0.338 | 0.238 |
| Cox and Snell R-square | 0.331 | 0.179 |
| Nagelkerke R-square | 0.476 | 0.318 |

*a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.*
*b. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.*

### 6.5.4. Classification Table

The predictive accuracy of a fitted logistic regression model can be done through an examination of the Classification Table. The Classification Table shows how far the predicted probabilities are in conformity with the actual outcomes. This table is obtained as a result of cross-classifying the outcome variable, in this case female work force participation, with a dichotomous variable whose values have been derived from the estimated logistic probabilities (Hosmer and Lemeshow, 2000 pp.156). The estimated probabilities are compared to a defined cut-off value, which is usually 0.5. For cases where estimated probabilities exceed 0.5 the derived dichotomous variable is coded as having value 1 or else having a value 0 *(ibid*).

**Table 6.9: Classification Table- Model I**

| Observed | | | Predicted | | | Observed | Predicted | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | FLFP | | Percentage Correct | | FLFP | | Percentage Correct |
| | | | 0 | 1 | | | 0 | 1 | |
| Step 0 a b | FLFP | 0 | 0 | 66 | 0.0 | Step 1b | 36 | 30 | 54.5 |
| | | 1 | 0 | 169 | 100.0 | | 14 | 155 | 91.7 |
| | Overall | | | | 71.9 | | | | 81.3 |

*a. Constant is included in the Model. b. The cut value is .500*

The Classification Table for Model I (Table 6.9.) shows that the model correctly predicts 81.3 percent observations as compared to the constant only model. Sensitivity which is defined as the proportion of observations with Y=1 that are correctly predicted by the model is 91.7 and specificity defined as the proportion of observations with Y=0 that are correctly predicted by the model is 54.5 percent. The estimated equation is 9.4 percentage points better at predicting responses than the constant only model (Rai and Mukherjee, 2018). The Classification Table for Model II (Table 6.10) shows that the model correctly predicts 88.1 percent observations as compared to the constant only model. Sensitivity is 20.6 and specificity is 99.5 percent. The estimated equation is 2.6 percentage points better at predicting responses than the constant only model (Rai and Mukherjee, 2018).

**Table 6.10: Classification Table for Model II**

| Observed | | | Predicted | | | Observed | Predicted | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | FLFP_PAID | | Percentage Correct | | FLFP_PAID | | Percentage Correct |
| | | | 0 | 1 | | | 0 | 1 | |
| Step 0 a b | FLFP_PAID | 0 | 201 | 0 | 100.0 | Step 1b | 200 | 1 | 99.5 |
| | | 1 | 34 | 0 | 0.0 | | 27 | 7 | 20.6 |
| | Overall | | | | 85.5 | | | | 88.1 |

*a. Constant is included in the Model. b. The cut value is .500*

### 6.5.5. Multicollinearity in the Model

One of the important problems encountered in linear and logistic regression is the problem of multicollinearity which may be defined as a statistical phenomenon in which two or more independent variables are highly correlated or associated (Midi, Sarkar and Rana, 2010). In the presence of multicollinearity, the estimates of the regression coefficients are unstable and may have large standard errors besides leading to inaccurate results (Afifi et. al., 2012, p. 143). It may also affect calculations regarding individual predictors, although the overall predictive power or reliability of the model is not reduced (Midi et. al., 2010). It may also

lead to incorrect signs and magnitudes of coefficients estimates, which result in erroneous conclusions about relationships between dependent and independent variables (*ibid).*

The detection of multicollinearity in logistic regression model is not straightforward as in the case of linear regression model. One of the most popular methods of detecting the presence of multicollinearity is to construct a correlation matrix of all the explanatory variables used in the analysis. The rule of thumb is that if the pair-wise or zero-order correlation coefficient between two regressors is high, usually greater than 0.8, then multicollinearity is a serious problem (Gujarati, 2004, p. 359). However, high zero-order correlations although sufficient, are not a necessary condition for detecting the presence of multicollinearity because it can exist even though the zero-order or simple correlations are comparatively low (say, less than 0.50) (*ibid*). Other information such as Variance Inflation Factor (VIF), Tolerance, Eigen values, Condition Index may also be used to detect the presence of multicollinearity. The Variance Inflation Factor (VIF) is defined as $VIF = \frac{1}{1-r_{ij}^2}$ where $r_{ij}$ is the partial correlation coefficient between two regressors. VIF shows how in the presence of multicollinearity the variance of an estimator is inflated. As $r_{ij}^2$ approaches 1, the VIF approaches infinity. That is, the variance of an estimator increases with increase in the degree of collinearity, and it can become infinite in the limit (*ibid*, p. 351). As a rule of thumb, if VIF exceeds 10 it indicates multicollinearity, but in weaker models, such as in logistic regression; values above 2.5 may be a cause for concern (Allison, 2001 cited in Midi et. al., 2010). The inverse of the VIF is the tolerance i.e. $\text{Tolerance} = \frac{1}{\text{VIF}}$. There is no formal cut off value to use tolerance for detecting multicollinearity. While Myers (1990) suggests a tolerance value below 0.1, Menard (2002) suggests that a tolerance value less than 0.2 to be an indication of the multicollinearity problem. As a rule of thumb, a tolerance of 0.1 or less can be a cause of concern (cited in *ibid*).

Sometimes Eigen values and condition indices are also used to detect the presence of multicollinearity. In the absence of collinearity, the Eigen values, condition indices or condition number will all equal unity, but with increase in collinearity, Eigen values will be both greater and smaller than unity. Multicollinearity problem is indicated by Eigen values close to zero and high condition indices (*ibid*). Another way of expressing these Eigen values is in the form of condition index which represent the square root of the ratio of the largest Eigen value to the Eigen value of interest (*ibid*). A large condition index indicates presence of multicollinearity. Belsley, Kuh and Welsch (1980) indicate that values of condition index greater than 30 can indicate serious problems (cited in Afifi et. al., 2012). An informal rule of

thumb is that multicollinearity is a cause of concern if the condition index is 15, and a cause of very serious concern if it is greater than 30, (Midi et. al., 2010).

**Table 6.11: Correlation Matrix of Explanatory Variables for Women's Work Participation**

| | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ | $X_8$ | $X_9$ | $X_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $X_1$ | 1 | | | | | | | | | |
| $X_2$ | 0.068 | 1 | | | | | | | | |
| $X_3$ | 0.177 | 0.051 | 1 | | | | | | | |
| $X_4$ | -0.010 | -0.068 | 0.108 | 1 | | | | | | |
| $X_5$ | 0.074 | -0.054 | -0.078 | -0.034 | 1 | | | | | |
| $X_6$ | -0.065 | 0.076 | 0.113 | 0.033 | -0.145 | 1 | | | | |
| $X_7$ | 0.173 | 0.024 | 0.008 | -0.025 | -0.070 | -0.062 | 1 | | | |
| $X_8$ | -0.430 | 0.028 | -0.344 | 0.176 | 0.047 | -0.031 | 0.053 | 1 | | |
| $X_9$ | -0.213 | 0.210 | -0.029 | -0.179 | -0.143 | 0.113 | 0.038 | 0.283 | 1 | |
| $X_{10}$ | 0.136 | 0.167 | -0.130 | -0.046 | 0.266 | 0.011 | 0.138 | -0.013 | -0.024 | 1 |

*Note: $X_1$-EDUCATION, $X_2$-FAM_STR, $X_3$-CHILD_06, $X_4$-MARITAL_STATUS, $X_5$-OCCUPATION_HEAD, $X_6$-MIGRANT, $X_7$-MPCE, $X_8$-AGE,$X_9$–AGE_SQU, $X_{10}$- LAND*

In the present analysis the pair wise correlation coefficients between the regressors exhibit values less than the cut off value of 0.8 as can be seen from Table 6.11. Further, for both the models as shown in Appendix B, the values for the Tolerance, VIF, Eigen Value and Condition Indices are well within the cut off range to rule out multicolleariy among the predictors (Rai and Mukherjee, 2018).

## 6.6. RESULTS AND DISCUSSION

The results of the binary logistic regression analyses which was conducted using SPSS software version 23 to determine the factors affecting female work participation in the study area is presented in the Tables 6.12 and 6.13 for both the models. The full model including all the predictor variables was tested against the constant only model and was found to be statistically significant, indicating that the set of predictors helped explain the work force participation behaviour of the women in the study area (Chi-square= 94.443 for Model I at $p$=0.000 at 8 degrees of freedom and Chi-square= 46.268 for Model II at $p$=0.000 at 9 degrees of freedom as shown in the Table 6.6 above). The pseudo R square values (Cox & Snell R Square=0.331, Nagelkerke R Square=0.476 and McFadden R square= 0.338) for Model I show that 33-48 percent of the variations in female work participation is explained by the set of predictor variables. The pseudo R square values for Model II are Cox & Snell R Square=0.179, Nagelkerke R Square=0.318 and McFadden R square= 0.238. The Classification Table shows the extent to which the predicted probabilities agree with the

actual outcomes. Here the percentage of correct predictions by the model is 81.3 percent for Model I and 88.1 percent for Model II (Rai and Mukherjee, 2018).

### 6.6.1. Results from the Logistic Regression

Tables 6.12 and 6.13 present the results of the binary logistic regression analyses. The *p*-values of the Wald statistics in Table 6.12 show that the variables that are significant in explaining female work force participation as per Model I are AGE, AGE_SQU, FAM_STR, and OCCUPATION_HEAD.

The estimated logit equation for Model I may therefore be written as:

**Predicted logit of FLFP = 1.850+ 0.104 AGE –0.006 AGE_SQU – 0.696 FAM_STR + 0.942 OCCUPATION_HEAD+ $u_i$**

Where $u_i$ is the stochastic error term**.**

**Table 6.12: Binomial Logistic Regression Estimates for Female Work Force Participation- Model I**

| **Variables in the Equation-Model I** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **B** | **S.E.** | **Wald** | **df** | **Sig.** | **Exp(B)** | **95% C.I.for EXP(B)** | |
| | | | | | | | | **Lower** | **Upper** |
| Step 1[a] | AGE | .104 | .021 | 24.884 | 1 | .000* | 1.109 | 1.065 | 1.155 |
| | AGE_SQU | -.006 | .001 | 23.323 | 1 | .000* | .994 | .991 | .996 |
| | EDUCATION | -.034 | .041 | .702 | 1 | .402 | .966 | .891 | 1.047 |
| | FAMILY_STR(1) | -.696 | .394 | 3.122 | 1 | .077*** | .498 | .230 | 1.079 |
| | CHILD_06(1) | -.511 | .496 | 1.062 | 1 | .303 | .600 | .227 | 1.586 |
| | MARITAL_STATUS(1) | .619 | .445 | 1.938 | 1 | .164 | 1.857 | .777 | 4.441 |
| | OCCUPATION_HEAD(1) | .942 | .385 | 5.980 | 1 | .014** | 2.566 | 1.206 | 5.459 |
| | MIGRANT(1) | -.358 | .447 | .639 | 1 | .424 | .699 | .291 | 1.681 |
| | Constant | 1.850 | .587 | 9.948 | 1 | .002** | 6.361 | | |

* significant at α=0.001,** significant at α=0.050 , *** significant at α=0.100
a.Variable(s) entered for step 1: AGE, AGE_SQU, EDUCATION, FAM_STR, CHILD_06, MARITAL_STATUS, OCCUPATION_HEAD, MIGRANT

The variables that are significant in explaining women's work participation in paid activities are EDUCATION, FAM_STR, CHILD_06, and LAND. The estimated logit equation for Model II may be written as:

**Predicted logit of FLFP_PAID = 0.353+ 0.121EDUCATION – 0.938FAM_STR – 2.165 CHILD_06 –0.755LAND+ $u_i$**

Where $u_i$ is the stochastic error term**.**

**Table 6.13: Binomial Logistic Regression Estimates for Female Work Force Participation- Model II**

| Variables in the Equation | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **B** | **S.E.** | **Wald** | **df** | **Sig.** | **Exp(B)** | **95% C.I.for EXP(B)** | |
| | | | | | | | | **Lower** | **Upper** |
| Step 1[a] | AGE | .001 | .027 | .001 | 1 | .973 | 1.001 | .950 | 1.055 |
| | AGE_SQ | -.003 | .002 | 1.928 | 1 | .165 | .997 | .993 | 1.001 |
| | EDUCATION | .121 | .048 | 6.362 | 1 | .012** | 1.129 | 1.027 | 1.241 |
| | FAMILY_STR(1) | -.938 | .458 | 4.185 | 1 | .041** | .392 | .159 | .961 |
| | CHILD_06(1) | -2.165 | 1.075 | 4.059 | 1 | .044** | .115 | .014 | .943 |
| | MARITAL_STATUS(1) | -.636 | .511 | 1.546 | 1 | .214 | .530 | .194 | 1.443 |
| | OCCUPATION_HEAD(1) | -.608 | .435 | 1.955 | 1 | .162 | .545 | .232 | 1.276 |
| | MIGRANT(1) | -.983 | .667 | 2.176 | 1 | .140 | .374 | .101 | 1.382 |
| | LAND | -.755 | .278 | 7.409 | 1 | .006** | .470 | .273 | .809 |
| | Constant | .353 | .687 | .263 | 1 | .608 | 1.423 | | |

** significant at α=0.050

a.Variable(s) entered on step 1: AGE, AGE_SQU, EDUCATION, FAM_STR, MARITAL_STATUS, OCCUPATION_HEAD, MIGRANT, LAND.

### 6.6.2. Interpretation of Log Odds and Odds Ratio

The interpretation of the results of the analysis involves two issues: (i) determination of the functional relationship between the outcome and the predictor variable, and (ii) determining the unit of change for the predictor variables. In linear regression the interpretation of the slope coefficients for the independent variables is straightforward as the slope coefficients represent the resulting change in the outcome variable for a unit change in the predictor variable. In the logistic regression model, the link function is the logit function and the slope coefficients represent the change in the logit corresponding to a change of one unit in the independent variable (Hosmer and Lemeshow, 2000 pp. 47-48). Therefore, the slope coefficient for a particular predictor variable (represented as the Bs in the results table) is the logarithm of the odds or the logit and shows the change in estimated log odds (or logit) of being in the work force i.e. for Y=1, for a unit change in that predictor variable, other predictors being held constant. The sign of the B-values (β) (showing the sign of the partial effects of each predictor) shows whether an explanatory variable has a positive or negative effect on the outcome variable, in this case female work force participation. For example, the estimated log of odds or logit, i.e. B value for the variable FAMILY_STR(1) is equal to minus (–) 0.696 for Model I, which implies that the estimated logit for a female in the workforce decreases by a factor 0.696 if the participant belongs to a joint family as compared to a participant in a nuclear family. The slope coefficients vary between plus and minus

infinity with a value 0 indicating that the given predictor variable does not affect the logit (Rai and Mukherjee, 2018).

According to Model I, the log of odds of a female participating in the workforce is significantly positively related to age (AGE) and the primary occupation of the household head (OCCUPATION_HEAD); and significantly inversely related to the age squared term (AGE_SQU) and structure of the household (FAM_STR). According to Model II, the log of odds of a female participating in paid work is significantly positively related to the years of schooling (EDUCATION); and significantly inversely related to family structure of the household (FAM_STR), the presence of a child below the age of six (CHILD_06), and the size of land owned by the household (LAND) (Rai and Mukherjee, 2018).

The parameter estimates of the binary logistic regression may also be easily interpreted in terms of the odds ratio which is nothing but the exponential of the B-values (β) and provides a directly understandable statistic for the relationship between the outcome variable and the specific predictor variable (given all the other predictor variables in the model are fixed) (Afifi et.al. pp. 275). For categorical predictor variables with two categories (i.e. X=0 and X=1), the odds ratio approximates how much more likely (or unlikely) it is for the outcome to be present among those with X=1 than among those with X=0 (Hosmer and Lemeshow, 2000 pp.49-50). While interpreting the odds ratio for a categorical variable it is important to keep in mind the coding for the two categories of the variable. In SPSS, the first or the last category can be chosen as the baseline category. Usually the absence of the factor is coded as 0 and the presence of the factor as 1. In the present analysis the lower category (i.e. X=0) is chosen as the baseline category. For a continuous variable X with slope coefficient β, the quantity exp (β) is interpreted as the ratio of the odds for a person with value (X+1) relative to the odds for a person with value X. Therefore, exp (β) is the incremental odds ratio which corresponds to an increase in one unit in variable X, assuming that the values of other X variables remain the same. The incremental odds ratio corresponding to the change of *k* units in X is exp ($k\beta$) (Afifi et. al., pp. 276). In SPSS, the odds ratio appears as Exp (B) in the "Variables in the Equation" table which may be interpreted in the following way. For example, the Exp (B) for FAM_STR(1) for Model I is 0.498, which implies that the odds of being in the workforce for a respondent from a joint family decreases by a factor 0.498 or 50.2 per cent, as compared to a respondent from a nuclear family. On the other hand, the odds ratio or Exp (B) value for EDUCATION in Model II is 1.129. This implies that with an increase in years of schooling of the respondent

by one year, the odds of the respondent being in the workforce increases by 12.9 per cent (Rai and Mukherjee, 2018).

The odds can also be expressed in percent terms which can be obtained by using the following expression: [{Exp(B)-1}*100]%. If the odds ratio is greater than 1, a unit increase in the predictor increases the odds in favour of the event (i.e. for Y=1) by the percent obtained. If the odds ratio is less than 1, a unit increase in the predictor reduces the odds by [1-{Exp(B) }*100]%. For example, the Exp (B) for FAM_STR(1) for Model I is 0.498 which implies that the odds of being in the work force for a respondent belonging to the joint family decreases by a factor 0.498 or by [{1-Exp(B)}*100]%= [{1-0.0.498}*100]%= 50.2 % as compared to being in the unitary family. On the other hand the odds ratio or Exp (B) value for EDUCATION in Model II is 1.129. This implies that with an increase in years of schooling of the respondent by 1 year, the odds of the respondent being in the work force increases by [{Exp(B)-1}*100]%= [{1.129-1}*100] %= 12.9 %.

**6.6.3. Interpretation of the Results of the Binary Logistic Regression**

Scholars have cited age as an important determinant of female work participation. The simultaneous demands made by children and work reduce women's work participation during periods of child-bearing and child-rearing as compared to women outside this age (Psacharopoulos and Tzannatos 1989). During the child-rearing period, female participation declines, but is expected to be the highest before the beginning of and a few years after the child-bearing period (Mon 2000). Reddy (1979) notes that although there is a clear cut negative relationship between female activity rates and child-bearing and child-rearing age-groups in the urban areas, there is no evidence of such an association in the rural areas. This may be due to the prevalence of the joint family system in the rural areas in which the older female members of the family assist in child-rearing (Reddy 1979) with older siblings also helping in the process. In Model I, AGE has a significant positive impact on female workforce participation whereas AGE_SQU has a negative significant effect, which shows the non-linear effect of age. Increase in age is associated with increased work participation up to a certain age, beyond which work participation decreases. This implies that younger women – who perform a greater share of household duties, as well as child-bearing and child-rearing activities – and older women are less likely to be in the workforce, as compared to middle-aged women. When we consider women's paid employment (Model II) however, the age of the respondent is not significant in explaining female work participation in the study areas.

Although theoretically there appears to be a positive correlation between female labour force participation and levels of education, empirical findings from developing countries present mixed results (Standing 1981 cited in Ackah *et al.* 2009). In some cases, education and female participation rates show only a marginal or non-linear relationship (Mon 2000). According to a study by Psacharopoulos and Tzannatos (1989), education has an ambiguous effect on women's participation in the labour force. They postulate that labour force participation rate is affected by the decision to participate in the labour market and by the decision of how much time to spend in the labour market. As regards the decision to participate in the labour market, education has a positive effect for two reasons. Firstly, if education is considered as an investment then the woman has to work in order to recover that cost of investment in human capital. Secondly, if education is considered a consumption activity, the woman will be induced to work due to higher earning potential as the opportunity cost of not working in terms of forgone earnings increases. As regards the duration of work, education has a positive effect as it raises the earning capacity and increases the cost of not working. On the other hand, higher earnings mean that the income target is reached earlier, allowing the woman to allocate a part of the higher earning to consume leisure and work less. The net effect of education on female labour force participation depends on which force dominates. Empirical studies have shown that female labor supply responds more to wage considerations (substitution effect) than to income, so that participation of educated females is higher than that for the less educated or uneducated (Psacharopoulos and Tzannatos, 1989).

In rural areas, non-farm paid jobs available to those with little or no education are mainly casual wage labour, where there is little association between the years of education and wage levels. However, education raises the reservation wage for these women through an increase in the productivity of time spent on their own farm and home production, which results in lower participation in wage/paid employment if the local labour market does not provide better opportunities (Unni, 1994). This implies that women with some education may prefer to remain outside the labour market altogether, preferably doing household work or working on family farms as unpaid family labour in the absence of remunerative non-farm employment opportunities. A negative association between the level of education and female labour force/work participation in paid activities may thus be postulated in rural areas.

In the present analysis, the coefficient of education as measured by the number of years of schooling is negative but does not significantly affect female work participation rates in Model I. Since work participation in Model I includes both paid and unpaid activities, it is

plausible that the level of education may weakly determine women's participation in the workforce. In Model II, however, it is positive and statistically significant. The value of Exp (B) for EDUCATION in Model II is 1.129; i.e. with a one-year increase in the number of years of schooling, the participation of women in paid employment increases by a factor of 1.129 or 12.9 per cent. An implication of this finding is that though the level of education has a non-significant effect on the work participation decisions, both paid and unpaid of women, the level of education of women who are in paid employment is higher than those who are not. This also implies that women with a higher level of education are employed in non-farm jobs, as employment in agriculture is primarily as unpaid family labour. Hypothesis (5) which states that education enhances women's participation in the labour market is accepted with some modification as education has been found to be positively associated with women's participation in paid work.

The kinship system and the joint family are still prevalent in rural India. In joint families with a large number of family members, a dichotomy is visible between men's work and women's work, with males being involved in paid activities and females in domestic activities. On the other hand, women of working age in joint families are assisted in their domestic activities and child care by older women and other female members of the household (Reddy 1979), which in turn increases their participation in paid activities or agricultural activities on the family farm. In the study area, family structure (FAM_STR) is a dummy variable with the variable taking a value of 1 if the respondent belongs to a joint family and 0 otherwise. The coefficient for this dummy is negative and statistically significant for both models, indicating that a respondent who belongs to a joint family as compared to a nuclear family was less likely to be in the workforce. In Model I, the odds of participating in work for a respondent from a joint family decreased by a factor of 0.498 or 50.2 per cent, whereas in Model II, the odds decreased by a factor of 0.392 or 60.8 per cent. If there is a single earning member in a nuclear family generally the husband, the wife is likely to work alongside him to supplement family labour on the farm or supplement family income through participation in paid activities. This explains the higher participation in the workforce of women belonging to nuclear families.

The presence of children may have a negative effect on women's participation in economic activities (Chaykowski and Powell 1999). Younger children especially, i.e. children below the age of six, may cause women to spend more time in child care while the presence of older children may reduce their work burden. Cohen (1970) observed that the presence of a child under the age of six was the most significant factor that determined labour

force participation of married women (cited in Anderson and Dimon 1998). In rural areas the presence of young children may not pose much of a problem for women's participation in agricultural activities, as older female children and female members of the household help with domestic work and child care while older male children assist in some agricultural activities. However, the spread of primary and secondary education has meant that school-going children cannot help in child care and household work as before.

As mentioned above, in nuclear families the presence of small children, particularly below the age of six, may hinder a woman's participation in economic activities. In the present study, the presence of children below the age of six was represented by the dummy variable CHILD_06, with the presence of one or more children below the age of six in the household being denoted by 1 and their absence by 0. In Model I, the presence of children under six years has a negative effect on female participation rates, but the results are not statistically significant. In Model II however, the variable has a significant negative effect. The results indicate that for women with children less than six years of age, the odds of participating in paid work decreased by a factor of 0.115 or 88.5 per cent. Since paid work involves working away from the vicinity of the household as opposed to unpaid work on family farms, taking care of young children may hinder women's participation in paid work.

The marital status of the respondent was another major influence on female labour force participation, as married women had larger household responsibilities than women who were not married (Mon 2000) which restricted their participation in the labour force. Being married influenced women's decision-making ability and also increased the value of non-market activities. In patriarchal family structures, women were expected to fulfill the role of mothers and home-makers, and men the role of breadwinners and heads of the household (Blau *et al.* 1998, p. 13 cited in Lisaniler and Bhatti 2005). Since such patriarchal family structures are widely prevalent in Indian society, marriage is expected to reduce the participation of women in labour market activities. We grouped the respondents into two categories: those currently married and those who were single/widowed/divorced/separated. The variable of marital status (MARITAL_STATUS) is a dummy with single/widowed/divorced/separated being the reference category. The results of the analysis show that for Model I, although the coefficient is positive, it is non-significant. For Model II the coefficient is negative, implying that women in the single/widowed/divorced/separated category participated more in paid employment but this is insignificant.

The decision of women to participate in the workforce is also influenced by the work status of their husbands. Women are more likely to work for cash if their husbands have no

source of income, and are more likely to work as unpaid family workers in the family business if their husbands are self-employed (Donahoe, 1999). Nam (1991) found that in households where the male head was self-employed in the tertiary sector, or was employed as a family worker, or was unemployed, married women's likelihood of participating in the labour market was two to three times more as compared to women from families with a higher social status controlling for age, number of children under 6 and marital status. Women from household where the heads were blue-collar wage workers had low female labour force participation (Nam, 1991). In our analysis, the primary occupation of the head of the household whether agricultural or non-agricultural, was considered a determinant of female work participation. The dummy variable (OCCUPATION_HEAD) took a value of 1 if the primary occupation was agriculture and 0 if it was non-agriculture. For Model I, the occupation of the head of the household is significant in explaining the work participation of the respondents, with the odds of a respondent participating in the workforce increasing by a factor of 2.566 if the occupation of the household head is agriculture. This implies that the odds of being in the workforce for women in agricultural households increases by almost three times, as against women in non-agricultural households. These results corroborate the findings outlined in the previous chapter regarding the significant role played by women in agricultural activities. For Model II, the variable OCCUPATION_HEAD although negative, does not have a significant effect on women's work participation.

Male migration has been significant in rural areas, particularly in the hill and mountain areas from where men have moved to lowland areas in search of better employment opportunities. This leads to women's heavier work burden and increases their drudgery as they now need to perform those tasks which were previously performed by men (Pande, 1996). A study of labour out-migration in households engaged in rice-farming in three districts of eastern Uttar Pradesh also reports an increase in the workload of women in nuclear households in the absence of males with women taking over many male-specific activities in rice farming. The study also notes that although women's decision-making capacity has increased due to migration of males, their work is hindered since they lack access to modern seed technology (Paris, Singh, Luis, and Hossain, 2005).

The impact of male migration on the labour market behaviour of women, however, is ambiguous. A theoretical model developed by Lokshin and Glinskaya (2009) predicts that male migration could have two effects on female labour market participation. First, as household income increases due to remittances it could lead to a decline in labour market participation of women. Secondly, women's productivity at home could increase or decrease

depending on the properties of the home production function, with effect on labour market participation being ambiguous. The overall effect on women's participation is therefore the result of the interaction of these factors.

To understand the effect of male migration on female work participation, the presence of a male migrant (MIGRANT) has been included in the present model as a dummy variable, with respondents in households with at least one male migrant being coded as 1 and households with no male migrants as 0. Though the presence of male migrants affects female work participation inversely in both models, it is not significant. The non-significant effect of MIGRANT in Model I is plausible, where both paid and unpaid activities of females have been considered. Irrespective of the presence of a male migrant women in rural households participate in paid as well as unpaid activities. In Model II women in households with at least one male migrant are less likely to participate in paid activities as household income may increase as a result of remittances (Lokshin and Glinskaya 2009). Women may also be forced to stay at home to perform household chores that were earlier performed by men, thus increasing their participation in unpaid domestic activities and leaving less time for participation in paid activities. This is contrary to the belief that male migration increases the participation of women in the workforce. Therefore hypothesis (4) that male out-migration has increased the work participation of hill women can thus be rejected in the present study.

Family income has been noted as an important determinant of female work participation. Nayyar (1987) writes that according to several scholars poverty has been regarded as "the single most important factor" which has an influence on participation rates for women and "cuts across regions, religions, age, and time." Low levels of earnings among males induce females to participate in economic activities to supplement family income, a phenomenon referred to as the "additional worker effect" (Reddy 1979). Alternatively, the participation of women in the labour market leads to an increase in total family income thereby postulating a positive relation between female labour force participation and total household income.

To avoid endogeneity in the present study, household income was approximated by using monthly per capita consumption expenditure (MPCE) as a proxy for family income. Compared to estimated income, estimated household expenditure is considered to be a better indicator of living standards particularly in household surveys conducted in developing countries (Mailu, Maritim, Yabann and Muhammed, n. d.). The univariate analysis for household MPCE was statistically insignificant in both models. This implies that women in rural areas irrespective of the level of expenditure or income participate in economic

activities, both paid and unpaid. Therefore, the hypothesis (6) that labour force/work participation of women is relatively higher in case of low income families is not supported by the results of the study.

Land is not only a vital asset in agricultural families but also an indicator of socio-economic status. Some micro studies have established a negative correlation between landlessness and female participation rates in rural areas in India. Given that landlessness is regarded as an indicator of poverty in rural areas, it appears logical that women in the landless category participate more in economic activities to supplement family income than women with land (Nayyar, 1987). Some studies however, find a positive relationship between women's work participation in agriculture and the size of landholding (Bhati and Singh, 1987). In our study we have measured the landholdings of households (LAND) in acres. The results according to the univariate analysis, indicate a non-significant relation between the size of landholding and female work participation and were thus not included in the logistic regression exercise. Since the study considers paid as well as unpaid work, land does not appear statistically significant as women in families with small or large holdings may be employed as unpaid labour on the family farm. Low prevalence of agricultural labour, and less stringent class and caste distinctions in the hill areas (relative to the plains) may also help explain the non-significant effect of land on women's paid and unpaid labour on family farms. In Model II however, the size of land owned by the family is statistically significant in explaining women's work participation in paid activities. The odds of a respondent participating in paid employment decreased by a factor of 0.470 or 53 per cent for every one-acre increase in land owned by the household. This can be attributed to the fact that smaller landholdings mean lower income from agriculture, inducing women to search for paid employment outside the household, in agriculture or non-agriculture (Rai and Mukherjee, 2018).

## 6.7. CONCLUSION

Our analysis shows that women's employment in the rural hill regions of West Bengal is characterised by a predominance of unpaid work as family labour in agriculture. While 71.9 per cent of working-age women in the study were employed in paid as well as unpaid employment only 14.5 per cent reported active participation in paid employment. This highlights the crucial role that women in the rural hill economy play through their involvement in unpaid farm employment and allied work which contributes significantly to household food and economic security.

The findings of the study show that the age of women has a non-linear effect on their participation in economic activities (paid as well as unpaid work). However, age does not show a significant effect on women's work decisions if we consider only paid work. This implies that younger women – on account of child-bearing and child-rearing activities and other household work – and older women may not participate in economic activities as much as middle-aged women.

The results of the study also indicate the significant positive effect of education on women's involvement in paid work. Women with higher levels of education preferred to take up jobs outside the agricultural sector, either in self-employment or wage employment. Self-employment in the region included petty trade, such as running a shop in the vicinity of the household, while wage employment included teaching, working as an ICDS helper, a mid-day meal cook or in a government office. Wage employment in agriculture was not common due to the prevalence of the labour exchange system of '*parma'*. Hypothesis (5) which states that education enhances women's participation in the labour market is accepted with some modification as education has been found to be positively associated with women's participation in paid work.

The structure of the family was also an important determinant of women's labour market behaviour, with women in nuclear families showing higher participation as compared to women in joint families. The presence of a single male breadwinner and the desire to augment family income in order to improve living standards may be contributing factors for higher work participation of women in nuclear families. This suggests a higher work burden for women in nuclear families where there is very little sharing of domestic responsibilities. The study also shows that the presence of children under the age of six lowered women's participation in paid work. The marital status of women was insignificant in determining women's involvement in economic activities in both models, although women's marital status and women's work participation had an inverse relation in the two models.

The presence of a male migrant in the family and the per capita consumption level of the household, a proxy for household income, had no significant effect on the work participation of women in the study. This is in contrast to studies that report higher female participation due to male out-migration. Therefore hypothesis (4) that male out-migration has increased the work participation of hill women can be rejected in the present study. The study also rejects hypothesis (6) that labour force/work participation of women is relatively higher in case of low income families since it is not supported by the results of the study. Finally, size of landholding was found to influence women's participation in paid activities in the

study area, with women in households with smaller landholdings showing a higher likelihood of participating in paid activities.

Since women in the rural hill region make significant contributions to family farms as unpaid labour, it is important to recognise their contributions through a proper valuation of their services. Recognising women as farmers and increasing their skills through training and education along with provision of extension services would help improve the position of women involved in farm activities. Diversification towards non-traditional agricultural activities such as horticulture, apiculture, pisciculture etc. would also help augment family income and improve rural well-being. Further, new jobs in the non-farm sector can increase the participation of women in paid work. Opportunities for self-employment in various farm and non-farm activities can also be explored such as processing of dairy products, pickling, production of jams and juices, handicrafts, and eco-tourism. Since higher levels of education induce women to take up paid work outside agriculture, increasing the level of women's education would lead to women's increased participation in paid work. The limiting effect of fertility on women's paid employment can be offset through provision of child-care facilities at work. Policy measures directed towards employment generation in the region should focus on women's employment for their empowerment and reduction in gender inequalities in the larger interest of the region and the nation.

**Chapter References**


- Ackah, C., Ahiadeke, C. and Fenny, A. P. (2009). *Determinants of female labour force participation in Ghana*. GDN Working Paper Series Working Paper No. 14.
- Actionaid. (2017). *Invisible work, invisible workers: The sub-economies of unpaid work and paid work: Action research on women's unpaid labour*. New Delhi: ActionAid Association.
- Afifi, A. May, S. and Clark, V. A. (2012). *Practical multivariate analysis* (Fifth Edition). London, New York: CRC Press, Taylor and Francis Group.
- Agarwal, B. (1989). Rural women, poverty and natural resources: Sustenance, stability and struggle for change. *Economic and Political Weekly*, *24* (43), WS46-WS65.
- Allison, P. (2012). *When you can safely ignore multicollinearity (*Blog post). Retrieved from. https://statisticalhorizons.com/multicollinearity viewed June 1, 2018
- Anderson, J. B., and Dimon, D. (1998). Married women's labour force participation in developing countries: The case of Mexico. *Estudios Económicos*, *13* (1(25)), 3–34.
- Atal, V. (2017). Say at home, or stay at home? Policy implications on female labour supply and empowerment. *Review of Economics of the Household*, *15* (4), 1081-1103.
- Bhati, J.P. and Singh, D.V. (1987). Women's contribution to agricultural economy in hill regions of north-west India. *Economic and Political Weekly*, *22* (17), WS7-WS11.
- Census of India. (2011). *Census of India, 2011*. Retrieved from http://www.censusindia.gov.in viewed on 12 October, 2017.
- Chakraborty, I. and Chakraborty, A. (2009). *Female work participation and gender differential in earning in West Bengal*. Kolkata: Institute of Development Studies.
- Chaykowski, R. P., and Powell, L. M. (1999). Women and the labour market: Recent trends and policy issues. *Canadian Public Policy/Analyse de Politiques* (*Supplement: Women in the Canadian Labour Market*), *Vol. 25*, November, S1–S25.
- Choudhary, N. and Parthasarathy, D. (2007). Gender, work and household food security. *Economic and Political Weekly*, *42* (6), 523-531.
- Cox, D. R. and Snell, E. J. (1989). *Analysis of binary data*. (2nd edition). London: Chapman and Hall.
- Donahoe, D. A. (1999). Measuring women's work in developing countries. *Population and Development Review*, *25* (3), 543–76.

- Government of India (GoI) (2013). *Key indicators of employment and unemployment in India, 2011–2012*: *NSS 68th Round*. National Sample Survey Office, Ministry of Programme Statistics and Implementation, Government of India, New Delhi.
- ------ (2014). *Household consumption of various goods and services in India 2011–12: NSS 68th Round*. National Sample Survey Office (NSSO), Ministry of Statistics and Programme Implementation, Government of India, New Delhi.
- Gujarati, D. N. (2004). *Basic econometrics* (Fourth Edition). New York: The McGraw-Hill Companies.
- Hensher, D. A., Rose, J. M., and Green, W. H. (2005). *Applied choice analysis: A Primer*. Cambridge: Cambridge University Press.
- Hosmer, D. W. and Lemeshow, S. (2002). *Applied logistic regression* (Second Edition). New York: John Wiley & Sons, Inc.
- International Labour Organisation (ILO). (2012). *Global employment trends for women*. Geneva: International Labour Office.
- Lisaniler F. G. and Bhatti, F. (2005). Determinants of female labour force participation: A study of North Cyprus. *Review of Social, Economic & Business Studies, 5/6,* 209 – 226.
- Lokshin, M. and Glinskaya, E. (2009). The effect of male migration on employment patterns of women in Nepal. *The World Bank Economic Review, 23* (3), 481–507.
- Mailu, S. K., Maritim, H. K., Yabann, W. K., and Muhammed, L. (n. d.). *Household income versus household expenditure: An illustration of the superiority of household expenditure as an estimator of living standards*. Kenya Agricultural Research Institute. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=A83811037466379117899971BEBF2B81?doi=10.1.1.197.791&rep=rep1&type=pdf, viewed July 20, 2018.
- Midi, H., Sarkar, S. K., and Rana, S. (2010). Collinearity diagnostics of binary logistic regression model. *Journal of Interdisciplinary Mathematics, 13* (3), 253–67.
- Mon, M. (2000). Determinants of female labour force participation in Burma: An empirical analysis of socio-economic survey data. *ABAC Journal, 21* (1).
- Nam, S. (1991). Determinants of female labor force participation: A study of Seoul, South Korea, 1970-1980. *Sociological Forum, 6* (4), 641-659.
- Nayyar, R. (1987). Female participation rates in rural India. *Economic and Political Weekly, 22* (51), 2207-2209+2211-2216.

- Pande, P. N. (1996). *Drudgery of the hill women.* New Delhi: Indus Publishing Company.
- Paris, T., Singh, A., Luis, J., and Hossain, M. (2005). Labour outmigration, livelihood of rice farming households, and women left behind: A case study in eastern Uttar Pradesh. *Economic and Political Weekly*, 40 (25), 2522–29.
- Park, H. A. (2013). An introduction to logistic regression: From basic concepts to interpretation with particular attention to nursing domain. *J Korean Acad Nurs., 43* (2), 154-164.
- Psacharopoulos, G. and Tzannatos, Z. (1989). Female labour force participation: An international perspective. *The World Bank Research Observer, 4*(2), 187-201.
- Rai, S. and Mukherjee, S. R. (2018). Covariates of rural female work participation: A study of the hill region of Darjeeling district in West Bengal, India. *Review of Agrarian Studies 1*(3).
- Reddy D. N. (1979). Female work participation in India: Facts, problems, and policies. *Indian Journal of Industrial Relations, 15* (2), 197-212.
- Sidh, S.N. and Basu, S. (2011). Women's contribution to household food and economic security: A study in the Garhwal Himalayas, India. *Mountain Research and Development 31*(2), 102-111.
- Sundar, P. (1981). Characteristics of female employment: Implications of research and policy. *Economic and Political Weekly, 16* (19), 863-871.
- United Nations Development Programme (UNDP). (2015). *Human development report, 2105: Work for human development*. New York, USA: United Nations Development Programme.
- Unni, J. (1994). Labour participation decisions of married women in rural India. *Indian Economic Review, 29* (2), 177–93.